日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

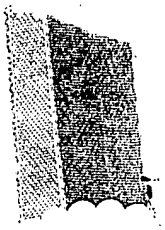
This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日
Date of Application: ２００４年　２月　６日

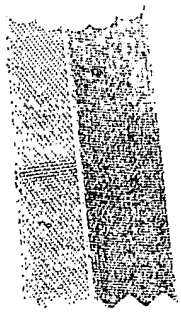
出願番号
Application Number: 特願２００４－０３１１５０

[ST. 10/C]: [ＪＰ２００４－０３１１５０]

出願人
Applicant(s): 株式会社日立製作所



CERTIFIED COPY OF
PRIORITY DOCUMENT



２００４年　４月２１日

特許庁長官
Commissioner,
Japan Patent Office

今井康夫

出証番号　出証特２００４－３０３４４９２

【書類名】　特許願
【整理番号】　340301787
【あて先】　特許庁長官殿
【国際特許分類】　G06F 03/06
【発明者】
　【住所又は居所】　神奈川県小田原市中里３２２番２号　株式会社日立製作所ＲＡＩＤシステム事業部内
　【氏名】　渡辺　治明
【発明者】
　【住所又は居所】　神奈川県小田原市中里３２２番２号　株式会社日立製作所ＲＡＩＤシステム事業部内
　【氏名】　占部　喜一郎
【発明者】
　【住所又は居所】　神奈川県小田原市中里３２２番２号　株式会社日立製作所ＲＡＩＤシステム事業部内
　【氏名】　山神　憲司
【特許出願人】
　【識別番号】　000005108
　【氏名又は名称】　株式会社日立製作所
【代理人】
　【識別番号】　100095371
　【弁理士】
　【氏名又は名称】　上村　輝之
【選任した代理人】
　【識別番号】　100089277
　【弁理士】
　【氏名又は名称】　宮川　長夫
【選任した代理人】
　【識別番号】　100104891
　【弁理士】
　【氏名又は名称】　中村　猛
【手数料の表示】
　【予納台帳番号】　043557
　【納付金額】　21,000円
【提出物件の目録】
　【物件名】　特許請求の範囲　1
　【物件名】　明細書　1
　【物件名】　図面　1
　【物件名】　要約書　1
　【包括委任状番号】　0110323

【書類名】特許請求の範囲
【請求項１】
ホスト端末に接続された記憶制御システムが有する記憶制御サブシステムにおいて、
データを論理的に記憶する論理記憶デバイスと、
前記論理記憶デバイスを有し、前記論理的に記憶されたデータを物理的に記憶する物理記憶デバイスと、
仮想記憶容量値が設定されることによって実装状態となる仮想記憶ユニットと、
前記設定された仮想記憶容量値を記憶するメモリと、
前記仮想記憶ユニットを認識している前記ホスト端末からリード要求又はライト要求を受信した場合に、前記仮想記憶ユニットにおける仮想記憶領域と、前記論理記憶デバイスにおける論理記憶領域とを対応付けて、前記仮想記憶領域を介して前記論理記憶領域と前記ホスト端末との間でデータをやり取りする記憶制御部と
を備え、
前記記憶制御部が、前記メモリに記憶された仮想記憶容量値を前記ホスト端末に通知し、前記ホスト端末において前記仮想記憶容量値が記憶された後、前記仮想記憶ユニットが前記ホスト端末に接続されている間は、前記通知した仮想記憶容量値が変更されないようにする、
記憶制御サブシステム。
【請求項２】
前記記憶制御サブシステムの保守のための処理を行う保守用端末が前記記憶制御部に接続されている場合において、
前記記憶制御部は、前記保守用端末又は前記保守用端末に接続された外部端末から、新たな前記仮想記憶ユニットを準備することのユニット準備要求を受け、前記ユニット準備要求に応答して、少なくとも前記仮想記憶容量値の記入欄を持ったグラフィカルユーザインターフェースを前記保守用端末又は前記外部端末に提供し、前記記入欄に入力された仮想記憶容量値を、前記設定された仮想記憶容量値として前記メモリに記憶させる、
請求項１記載の記憶制御サブシステム。
【請求項３】
前記記憶制御サブシステムが、２つの記憶ユニットから成るユニットペアを形成し、一方の記憶ユニットをプライマリ記憶ユニット、他方の記憶ユニットをセカンダリ記憶ユニットとして、プライマリ記憶ユニット内のデータをセカンダリ記憶ユニットにコピーするスナップショットを行うようになっている場合、
前記物理記憶デバイスには複数の前記論理記憶デバイスが設けられており、
前記複数の論理記憶デバイスには、前記仮想記憶領域に対応付けられ得る論理記憶領域を持った２以上の第１論理記憶デバイスと、前記仮想記憶領域に対応付けられることのない論理記憶領域を持った１以上の第２論理記憶デバイスとが含まれており、
前記１以上の第２論理記憶デバイスが、前記ホスト端末に接続される１つのリアル記憶ユニットを構成し、
前記記憶制御部は、前記リアル記憶ユニットを前記プライマリ記憶ユニットとし、前記仮想記憶ユニットをセカンダリ記憶ユニットとしたユニットペアを形成して前記スナップショットを行う、
請求項１記載の記憶制御サブシステム。
【請求項４】
前記記憶制御部は、前記仮想記憶容量値を前記ホスト端末に通知していない場合に、前記仮想記憶ユニットと前記リアル記憶ユニットの前記ユニットペアを形成するならば、前記リアル記憶ユニットの記憶容量値と同じ値を前記仮想記憶ユニットの記憶容量値として前記ホスト端末に通知する、
請求項３記載の記憶制御サブシステム。
【請求項５】
前記記憶制御部は、前記仮想記憶ユニットの相手となる前記リアル記憶ユニットが見つ

からない場合に、前記ホスト端末から前記仮想記憶ユニットに対してリード要求又はライト要求を受けたならば、前記仮想記憶ユニットが未実装状態であると前記ホスト端末に通知し、その後、前記相手となるリアル記憶ユニットが見つかったならば、前記リアル記憶ユニットの記憶容量値と同じ値を前記仮想記憶ユニットの記憶容量値として前記ホスト端末に通知する、
請求項４記載の記憶制御サブシステム。
【請求項６】
前記記憶制御部は、前記ホスト端末から所定コマンドを受信した場合に、前記メモリに記憶されている仮想記憶容量値を前記ホスト端末に通知する、
請求項１記載の記憶制御サブシステム。
【請求項７】
前記記憶制御サブシステムの保守のための処理を行う保守用端末が前記記憶制御部に接続されており、且つ、前記ホスト端末が行う所定の処理で、前記ホスト端末が、前記記憶した仮想記憶容量値を消去するようになっている場合において、
前記記憶制御部は、前記仮想記憶容量値を前記ホスト端末に通知した後に、前記保守用端末又は前記保守用端末に接続された外部端末から、前記通知した仮想記憶容量値の更新要求を受けたならば、前記ホスト端末と前記仮想記憶ユニットとが接続されていない間に、前記保守用端末又は前記外部端末から新たな仮想記憶容量値を受けて前記メモリに記憶させ、且つ、前記ホスト端末に前記所定の処理を行わせることで、前記ホストに記憶されている旧い前記仮想記憶容量値が消去されるようにした後、前記メモリに記憶させた新たな仮想記憶容量値を前記ホスト端末に通知する、
請求項１記載の記憶制御サブシステム。

【書類名】明細書
【発明の名称】仮想記憶ユニットを備えた記憶制御サブシステム
【技術分野】
【０００１】
本発明は、仮想記憶ユニットを備えた記憶制御サブシステムに関する。
【背景技術】
【０００２】
例えば、大容量のデータを取り扱う基幹業務用の記憶システムでは、ホストコンピュータ（以下、「ホスト端末」と言う）とは、別体に構成された記憶制御サブシステムを用いてデータが管理されている。この記憶制御サブシステムは、例えば、多数のディスク型記憶装置をアレイ状に配置して構成されているＲＡＩＤ（Redundant Array of Independent Inexpensive Disks）である。
【０００３】
例えば、特開平９－２８８５４７号公報には、直列に並べられた複数の実デバイスを１つの仮想デバイスとして提供することが開示されている（段落２３）。
【０００４】
また、例えば、特開平１１－３４５１５８号公報には、正副ボリュームへの更新を各ボリュームの更新情報にそれぞれ記録する正副１対のボリュームに対する差分ダンプ取得方法であって、ミラー状態移行時に前記副ボリュームの更新情報を前記正ボリュームの更新情報に一致させ、ダンプ開始時に前記ミラー状態からミラー一時解除状態へ移行して前記正ボリュームの更新情報を初期化するとともに前記副ボリュームから更新部分のダンプを得ることが開示されている。
【０００５】
【特許文献１】特開平９－２８８５４７号公報。
【特許文献２】特開平１１－３４５１５８号公報。
【発明の開示】
【発明が解決しようとする課題】
【０００６】
ところで、例えば、記憶制御サブシステムは、１又は複数のディスク型記憶装置上に用意される複数の論理的な記憶デバイス（以下、Logical Deviceを略して「ＬＤＥＶ」と言う）のうち、１以上のＬＤＥＶから構成される１つの記憶ユニットをＬＵ（Logical Unit）としてホスト端末に提供するものがある。このような記憶制御システムでは、例えば、１つのＬＵをプライマリＬＵ、別の１つのＬＵをセカンダリＬＵとしたＬＵペアが形成されて、プライマリＬＵ内の更新前データをセカンダリＬＵにコピーすること（すなわち、いわゆるスナップショットを作成すること）が行われる場合がある。
【０００７】
一般に、ＬＵペアを構成する双方のＬＵの記憶容量は同じである。このため、プライマリＬＵの記憶容量が大きいと、セカンダリＬＵとしても記憶容量の大きいＬＵが必要となる。しかし、上述したスナップショットでは、コピーされる更新前データのデータサイズが小さい場合がある。そのため、セカンダリＬＵの記憶容量は無駄に大きいことになってしまう場合がある。
【０００８】
また、例えば、記憶制御サブシステムは、ホスト端末からリードキャパシティコマンドを受けた場合、その記憶制御サブシステムが有するＬＵの記憶容量をホスト端末に通知することがある。ここで、記憶制御サブシステムが、同一のＬＵについて、過去に通知した記憶容量と別の記憶容量を通知すると、ホスト端末に混乱を生じさせてしまう可能性がある。
【０００９】
従って、本発明は、例えば、少なくとも次のいずれかの目的を達成するものである。
(１)スナップショット作成のために、記憶容量の大きいＬＵペアを形成しても、無駄な空

き記憶領域を少なくする記憶制御サブシステムを提供する。
(２)ホスト端末に混乱を生じさせてしまう可能性を低減する記憶制御サブシステムを提供する。

【００１０】

本発明の他の目的は、後述の説明から明らかになるであろう。

【課題を解決するための手段】

【００１１】

本発明に従う記憶制御サブシステムは、ホスト端末に接続された記憶制御システムが有する記憶制御サブシステムであって、データを論理的に記憶する論理記憶デバイスと、前記論理記憶デバイスを有し、前記論理的に記憶されたデータを物理的に記憶する物理記憶デバイスと、仮想記憶容量値が設定されることによって実装状態となる仮想記憶ユニットと、前記設定された仮想記憶容量値を記憶するメモリと、記憶制御部とを備える。記憶制御部は、前記仮想記憶ユニットを認識している前記ホスト端末からリード要求又はライト要求を受信した場合に、前記仮想記憶ユニットにおける仮想記憶領域と、前記論理記憶デバイスにおける論理記憶領域とを対応付けて（例えば動的に対応付けて）、前記仮想記憶領域を介して前記論理記憶領域と前記ホスト端末との間でデータをやり取りする。前記記憶制御部は、前記メモリに記憶された仮想記憶容量値を前記ホスト端末に通知し、前記ホスト端末において前記仮想記憶容量値が記憶された後、前記仮想記憶ユニットが前記ホスト端末に接続されている間は、前記通知した仮想記憶容量値が変更されないようにする。

【００１２】

本発明に従う記憶制御方法は、例えば、データを論理的に記憶する論理記憶デバイスと、前記論理記憶デバイスを有し、前記論理的に記憶されたデータを物理的に記憶する物理記憶デバイスと、仮想記憶容量値が設定されることによって実装状態となる仮想記憶ユニットと、前記設定された仮想記憶容量値を記憶するメモリとを備える記憶制御サブシステムにおいて行われる方法である。前記記憶制御方法は、前記仮想記憶ユニットを認識している前記ホスト端末からリード要求又はライト要求を受信した場合に、前記仮想記憶ユニットにおける仮想記憶領域と、前記論理記憶デバイスにおける論理記憶領域とを対応付けるステップと、前記仮想記憶領域を介して前記論理記憶領域と前記ホスト端末との間でデータをやり取りするステップと、前記メモリに記憶された仮想記憶容量値を前記ホスト端末に通知するステップと、前記通知するステップを実行し、前記ホスト端末において前記仮想記憶容量値が記憶された後、前記仮想記憶ユニットが前記ホスト端末に接続されている間は、前記通知した仮想記憶容量値が変更されないようにするステップとを有する。

【発明の効果】

【００１３】

本発明によれば、ホスト端末に混乱を生じさせてしまう可能性を低減する記憶制御サブシステムが提供される。

【発明を実施するための最良の形態】

【００１４】

以下、図面を参照して本発明の一実施形態を説明する。

【００１５】

図１は、本発明の一実施形態に係る記憶制御システムの外観の概略を示す。

【００１６】

記憶制御システム６００は、例えば、基本筐体１０と複数の増設筐体１２とから構成することができる（基本筐体１１のみで構成されても良い）。

【００１７】

基本筐体１０は、記憶制御システム６００の最小構成単位である。この基本筐体１０には、例えば、複数のディスク型記憶装置（例えばハードディスクドライブ（ＨＤＤ））３００と、複数の制御パッケージ（例えば後述するチャネル制御部又はディスク制御部）１０５と、複数の電源ユニット４００と、複数のバッテリユニット５００とがそれぞれ着脱可能に設けられている。また、基本筐体１０には、複数の冷却ファン１３が設けられてい

る。

【００１８】

各増設筐体１２は、記憶制御システム６００のオプションであり、例えば、１つの基本筐体１０に最大４個の増設筐体１２を接続することができる。各増設筐体１２には、複数の冷却ファン１３が設けられている。また、各増設筐体１２には、複数のディスク型記憶装置３００と、複数の電源ユニット４００と、複数のバッテリユニット５００とがそれぞれ着脱可能に設けられており、それらの各々は、例えば、基本筐体１０に設けられた制御パッケージ１０５が有する制御機能により制御される。

【００１９】

図２は、本発明の一実施形態に係る記憶システムの全体構成例を示す。

【００２０】

この記憶システム１の基本的な構成要素は、１又は複数のホスト端末２００Ａ～２００Ｄと、記憶制御システム６００である。

【００２１】

ホスト端末（上位装置）２００Ａ～２００Ｄの各々は、例えば、ＣＰＵ（Central Processing Unit）、不揮発性及び／又は揮発性のメモリ（例えばROM又はRAM）、及びハードディスク等をハードウェア資源として備えたコンピュータシステム（例えば、パーソナルコンピュータ又はワークステーション）である。各ホスト端末２００Ａ～２００ＤのＣＰＵが、メモリに格納された各種コンピュータプログラムを読込んで実行することにより、コンピュータプログラムとハードウェア資源（例えばメモリ）とが協働した処理が行われて、種々の機能が実現される。ホスト端末２００Ａ～２００Ｄの各々は、種々の方法で記憶制御システム６００に接続することができる。

【００２２】

例えば、ホスト端末２００Ａ～２００Ｂは、第一通信ネットワーク（例えば、ＬＡＮ、インターネット又は専用回線、以下、ＬＡＮであるとする）８２０を介して記憶制御システム６００に接続されている。ＬＡＮ８２０を介して行われるホスト端末２００Ａ及び２００Ｂと記憶制御システム６００との間の通信は、例えばＴＣＰ／ＩＰプロトコルに従って行われる。ホスト端末２００Ａ及び２００Ｂから記憶制御システム６００に対して、ファイル名指定によるデータアクセス要求（ファイル単位でのデータ入出力要求、以下、「ファイルアクセス要求」と言う）が送信される。

【００２３】

また、例えば、ホスト端末２００Ｂ及び２００Ｃは、第二通信ネットワーク（例えば、ＳＡＮ（Storage Area Network）、以下、ＳＡＮであるとする）８２１を介して記憶制御システム６００に接続されている。ＳＡＮ８２１を介して行われるホスト端末２００Ｂ及び２００Ｃと記憶制御システム６００との間の通信は、例えばファイバチャネルプロトコルに従って行われる。ホスト端末２００Ｂ及び２００Ｃから記憶制御システム６００に対して、例えば、ブロック単位のデータアクセス要求（以下、「ブロックアクセス要求」と言う）が送信される（なお、ブロック単位とは、後述のディスク型記憶装置３００上の記憶領域におけるデータの管理単位であるブロックを単位としたものである）。

【００２４】

また、例えば、ホスト端末２００Ｄは、ＬＡＮ８２０やＳＡＮ８２１等のネットワークを介さずに記憶制御システム６００に接続されている。ホスト端末２００Ｄは、例えば、メインフレームコンピュータとすることができる。ホスト端末２００Ｄと記憶制御システム６００との間の通信は、例えば、FICON（Fibre Connection：登録商標）、ESCON（Enterprise System Connection：登録商標）、ACONARC（Advanced Connection Architecture：登録商標）、FIBARC（Fibre Connection Architecture：登録商標）等の通信プロトコルに従って行われる。ホスト端末２００Ｄから記憶制御システム６００に対して、例えば、これらの通信プロトコルのうちのいずれかに従ってブロックアクセス要求が送信される。

【００２５】

ＬＡＮ８２０及びＳＡＮ８２１の少なくとも一方には、例えば、バックアップ記憶制御システム９１０が接続されている。バックアップ記憶制御システム９１０は、例えば、１又は複数個のディスク系デバイス（例えば、ＭＯ、ＣＤ－Ｒ、又はＤＶＤ－ＲＡＭ）から選択されたディスク系デバイスにデータを格納する記憶制御システムであっても良いし、後に詳述する記憶制御システム６００のような記憶制御システムであっても良いし、１又は複数個のテープ系デバイス（例えば、ＤＡＴテープ、カセットテープ、オープンテープ又はカートリッジテープ）から選択されたテープ系デバイスにデータを格納する記憶制御システムであっても良い。バックアップ記憶制御システム９１０は、例えば、ＬＡＮ８２０或いはＳＡＮ８２１（又は、更に、バックアップ制御システム９１０に接続されているホスト端末）を介して、記憶制御システム６００に記憶されているデータを受信し、自分が備えている記憶装置（例えばテープ系デバイス）にそのデータを格納する。

【００２６】

また、ＬＡＮ８２０及びＳＡＮ８２１のうちの少なくともＬＡＮ８２０には、例えば、管理サーバ８１９が接続されていても良い。管理サーバ８１９は、例えば、ホスト端末と別のホスト端末との間の通信や、ホスト端末と記憶制御システム６００との間の通信の中継を行っても良い。

【００２７】

また、ホスト端末２００Ａ～２００Ｄは、第三通信ネットワーク（例えばＬＡＮ）を介して相互に接続されていても良い。

【００２８】

記憶制御システム６００は、例えば、ＲＡＩＤ（（Redundant Array of Independent Inexpensive Disks））システムである。記憶制御システム６００は、ホスト端末２００Ａ～２００Ｄから受信したコマンドに従う制御を行う。記憶制御システム６００は、記憶制御サブシステム（ディスクアレイ装置）１０２と、保守用端末（以下、Service Processorを略して「ＳＶＰ」と記載）１６０とを備える。記憶制御サブシステム１０２は、記憶制御装置１００と、記憶装置ユニット１０１とを備える。記憶制御装置１００は、１又は複数のチャネル制御部１１０Ａ～１１０Ｄと、１又は複数のキャッシュメモリ（以下、Cache Memoryを略して「ＣＭ」と記載）１３０と、１又は複数の共有メモリ（以下、Shared Memoryを略して「ＳＭ」と記載）１２０と、１又は複数のディスク制御部１４０Ａ～１４０Ｄと、接続部１５０とを備える。記憶装置ユニット１０１は、１以上の物理ディスク群５を備える。１以上の物理ディスク群５の各々は、アレイ状に配列された複数のディスク型記憶装置３００を有する。

【００２９】

チャネル制御部１１０Ａ～１１０Ｄの各々は、ハードウェア回路、ソフトウェア、又はそれらの組み合わせで構成することができる。各チャネル制御部１１０Ａ～１１０Ｄは、記憶制御装置１００（例えば基本筐体１０）に対して着脱可能であり、チャネルアダプタと呼ばれることがある。各チャネル制御部１１０Ａ～１１０Ｄは、例えば、プロセッサやメモリ等が実装されたプリント基板と、メモリに格納された制御プログラムとをそれぞれ備えており、これらのハードウェアとソフトウェアとの協働作業によって、所定の機能を実現する。チャネル制御部１１０Ａ～１１０Ｄの各々は、多重化（例えば二重化）されており、１つのチャネル制御部が破損しても他のチャネル制御部で動作するようになっている。チャネル制御部１１０Ａ～１１０Ｄは、ＳＭ１２０内の制御情報等（例えば後述のＬＵ－ＬＤＥＶ管理テーブル）を参照しつつ、ホスト端末から受信したコマンドに従う処理を実行する。チャネル制御部１１０Ａ～１１０Ｄのうち、チャネル制御部１１０Ｃを例に採り、ディスク制御部１４０Ａ～１４０Ｄの動作も含めて先に説明すると、例えば、チャネル制御部１１０Ｃは、ホスト端末２００Ａ又は２００Ｂから、リード要求を含んだＩ／Ｏ要求（入／出力要求、ここではブロックアクセス要求）を受信すると、読出しコマンドをＳＭ１２０に記憶させると共に、ＣＭ１３０にキャッシュ領域を確保する。ディスク制御部１４０Ａ～１４０Ｄは、ＳＭ１２０を随時参照しており、未処理の読出しコマンドを発見すると、ディスク型記憶装置３００からデータ（典型的には、ホスト端末２００Ａ～

２００Ｄとディスク型記憶装置３００との間でやり取りされるユーザデータ）を読み出して、ＣＭ１３０に確保された上記キャッシュ領域に記憶させる。チャネル制御部１１０Ｃは、ＣＭ１３０に移されたデータをキャッシュ領域から読み出し、そのデータをリード要求発行元のホスト端末２００Ａ又は２００Ｂに送信する。

【００３０】

また、例えば、チャネル制御部１１０Ｃは、ホスト端末２００Ｂ又は２００Ｃから、ライト要求を含んだＩ／Ｏ要求を受信すると、書込みコマンドをＳＭ１２０に記憶させると共に、ＣＭ１３０にキャッシュ領域を確保し、受信したＩ／Ｏ要求に含まれているデータを、上記確保したキャッシュ領域に記憶させる。その後、チャネル制御部１１０Ｃは、ライト要求発行元のホスト端末２００Ｂ又は２００Ｃに対して書込み完了を報告する。そして、ディスク制御部１４０Ａ～１４０Ｄは、ＳＭ１２０を随時参照しており、未処理の書込みコマンドを発見すると、その書込みコマンドに従って、ＣＭ１３０に確保された上記キャッシュ領域からデータを読出し、そのデータを所定のディスク型記憶装置３００に記憶させる。

【００３１】

上記の処理は、他のチャネル制御部１１０Ａ～１１０Ｂ及び１１０Ｄも行うことができる。なお、チャネル制御部１１０Ａ～１１０Ｂは、ファイルアクセス要求をブロックアクセス要求に変換した後に（例えば、自分が持っているファイルシステムに従い、ファイル要求に含まれているファイル名を論理ブロックアドレスに変換した後に）、上記の処理を行う。

【００３２】

ディスク制御部１４０Ａ～１４０Ｄは、ハードウェア回路、ソフトウェア、又はそれらの組み合わせで構成することができる。各チャネル制御部１４０Ａ～１４０Ｄは、記憶制御装置１００（例えば基本筐体１０又は増設筐体１２）に対して着脱可能であり、ディスクアダプタと呼ばれることがある。ディスク制御部１４０Ａ～１４０Ｄは、例えば、プロセッサやメモリ等が実装されたプリント基板と、メモリに格納された制御プログラムとをそれぞれ備えており、これらのハードウェアとソフトウェアとの協働作業によって、所定の機能を実現する。ディスク制御部１４０Ａ～１４０Ｄの各々は、多重化（例えば二重化）されており、１つのディスク制御部が破損しても他のディスク制御部で動作するようになっている。ディスク制御部１４０Ａ～１４０Ｄは、ＳＭ１２０内の制御情報等（例えば後述のＬＵ－ＬＤＥＶ管理テーブル）を参照しつつ、各物理ディスク群５に含まれる各ディスク型記憶装置３００との間のデータ通信を制御するものである。各ディスク制御部１４０Ａ～１４０Ｄと各ディスク型記憶装置３００とは、例えば、ＳＡＮ等の通信ネットワークを介して接続されており、ファイバチャネルプロトコルに従ってブロック単位のデータ転送を行う。また、ディスク制御部１４０Ａ～１４０Ｄは、ディスク型記憶装置３００の状態を随時監視しており、この監視結果は内部の通信ネットワーク（例えばＬＡＮ）１５１を介してＳＶＰ１６０に送信される。

【００３３】

１又は複数のＣＭ１３０は、例えば、揮発性又は不揮発性のメモリである。ＣＭ１３０には、キャッシュ領域が確保され、そこに、チャネル制御部１１０Ａ～１１０Ｄとディスク制御部１４０Ａ～１４０Ｄとの間で送受されるデータが記憶される。そのデータは、複数のＣＭ１３０により多重管理されても良い。

【００３４】

１又は複数のＳＭ１２０は、例えば不揮発性のメモリから構成され、制御情報等を記憶する（例えば、制御情報等は、複数のＳＭ１２０により多重管理されても良い）。制御情報等には、例えば、チャネル制御部１１０Ａ～１１０Ｄとディスク制御部１４０Ａ～１４０Ｄとの間でやり取りされる種々のコマンドや、キャッシュ管理テーブルや、ディスク管理テーブルや、ＬＵ－ＬＤＥＶ管理テーブルがある。キャッシュ管理テーブルは、例えば、キャッシュ領域と、後述するＬＤＥＶの論理アドレスとの対応関係が書かれたテーブルである。ディスク管理テーブルは、各ディスク型記憶装置３００を管理するためのテーブ

ルであり、例えば、各ディスク型記憶装置３００毎に、ディスクＩＤ、ベンダ、記憶容量、ＲＡＩＤレベル、使用状況（例えば使用中か未使用か）等を有する。ＬＵ－ＬＤＥＶ管理テーブルは、後述するＬＤＥＶを管理するためのテーブルであり、例えば、各ＬＤＥＶ毎に、論理パス情報（例えばポート番号、ターゲットＩＤ及びＬＵＮ）、アドレス管理情報（例えば、ディスク型記憶装置３００上の物理アドレスとＬＤＥＶにおける論理アドレスとの対応関係）、記憶容量及びＲＡＩＤレベルを有する。なお、物理アドレスとは、例えば、ディスク型記憶装置３００のＩＤ、ディスクヘッド番号、及びセクタ数を含んだアドレス情報である。論理アドレスとは、例えば、ＬＵＮ（Logica Unit Number）、ＬＤＥＶ番号、及び論理ブロックアドレスを含んだアドレス情報である。

【００３５】

接続部１５０は、各チャネル制御部１１０Ａ～１１０Ｄと、各ディスク制御部１４０Ａ～１４０Ｄと、ＣＭ１３０と、ＳＭ１２０とを相互に接続するものである。チャネル制御部１１０Ａ～１１０Ｄ、ＣＭ１３０、ＳＭ１２０及びディスク制御部１４０Ａ～１４０Ｄ間でのデータやコマンドの授受は、接続部１５０を介することにより行われる。接続部１５０は、例えば、ユーザデータが通過する第１サブ接続部と、制御情報等が通過する第２接続部とを含んでいる。第１サブ接続部には、各チャネル制御部１１０Ａ～１１０Ｄ、各ディスク制御部１４０Ａ～１４０Ｄ及びＣＭ１３０が接続され、第２サブ接続部には、各チャネル制御部１１０Ａ～１１０Ｄ、各ディスク制御部１４０Ａ～１４０Ｄ及びＳＭ１２０が接続される。第１サブ接続部及び第２サブ接続部のうちの少なくとも第１サブ接続部は、例えば、高速スイッチングによりデータ伝送を行う超高速クロスバスイッチ等の高速バスである。

【００３６】

複数のディスク型記憶装置３００の各々は、例えば、ハードディスクドライブ或いは半導体メモリ装置等である。複数のディスク型記憶装置３００のうちの２以上の所定数のディスク型記憶装置３００によって、ＲＡＩＤグループ２が構成されている。ＲＡＩＤグループ２は、例えば、パリティグループ又はエラーコレクショングループとも呼ばれることがあり、ＲＡＩＤの原理に従ったディスク型記憶装置３００のグループである。同じＲＡＩＤグループ２に属する２以上のディスク型記憶装置３００は、例えば異なるマザーボード上に搭載され、一つのディスク型記憶装置３００が故障しても、残りの他のディスク型記憶装置３００のデータを用いて、その故障したディスク型記憶装置３００のデータを復元できるように構成されている。このＲＡＩＤグループ２の提供する物理的な記憶領域上に、論理的な記憶デバイスである複数のＬＤＥＶ（Logical Device）が設定され、複数のＬＤＥＶのうちの１以上のＬＤＥＶが、ＬＵＮ（Logical Unit Number）を持った１つのＬＵ（Logical Unit）３１０として、記憶制御装置１００からホスト端末２００Ａ～２００Ｄに提供される。各ＬＵ３１０は、例えば、プライマリＬＵ（データコピー元ＬＵ）として、セカンダリＬＵ（データコピー先ＬＵ）である別ＬＵ３１０とペアになる場合があり、その場合、そのＬＵ３１０内の全部又は一部のデータ（例えば更新前のデータ）が、別ＬＵ３１０にコピーされることがある。また、各ＬＵ３１０は、例えば、セカンダリＬＵとして、プライマリＬＵである他のＬＵ３１０とペアになる場合があり、その場合、別ＬＵ３１０内の全部又は一部のデータ（例えば更新前のデータ）が、そのＬＵ３１０にコピーされることがある。

【００３７】

ＳＶＰ１６０は、ストレージシステム６００を保守又は管理するためのコンピュータマシンである。ＳＶＰ１６０は、例えば、内部ＬＡＮ等の通信ネットワーク１５１を介して、記憶制御システム６００の各構成要素（例えば、各チャネル制御部１１０Ａ～１１０Ｄ及び各ディスク制御部１４０Ａ～１４０Ｄ）から情報を収集することができる。具体的には、例えば、記憶制御システム６００の各構成要素（例えばチャネル制御部又はディスク制御部）に搭載されているＯＳ（オペレーティングシステム）、アプリケーションプログラム、ドライバソフトウェア等が、その構成要素で発生した障害発生に関する障害発生情報を出力するようになっていて、ＳＶＰ１６０が、その障害発生情報を受信することがで

きる。ＳＶＰ１６０が受ける情報としては、例えば、装置構成、電源アラーム、温度アラーム、入出力速度（例えば、記憶制御装置１００が一定時間当たりに受信したＩ／Ｏ要求の数）等がある。また、例えば、ＳＶＰ１６０は、オペレータによる操作に応答して、例えば、ディスク型記憶装置３００の設定や、ＬＤＥＶの設定や、チャネル制御部１１０Ａ～１１０Ｄにおいて実行されるマイクロプログラムのインストール等を行うことができる。また、ＳＶＰ１６０は、例えば、記憶制御システム６００の動作状態の確認や故障部位の特定、チャネル制御部１１０で実行されるオペレーティングシステムのインストール等の作業を行うこともできる。また、例えば、ＳＶＰ１６０は、ＬＡＮや電話回線等の通信ネットワークを介して外部保守センタ（図示せず）と接続されていて、その外部保守センタに、記憶制御システム６００の各構成要素から受信した障害発生情報等を通知しても良い。また、ＳＶＰ１６０は、記憶制御システム６００に内蔵されている形態とすることもできるし、外付けされている形態とすることもできる。

【００３８】

以上が、記憶制御システム６００についての基本的な説明である。なお、この記憶制御システム６００では、例えば、１つのチャネル制御部と１つのディスク制御部とが１つのモジュールとして一体的に構成されて、その１つのモジュールで、チャネル制御部及びディスク制御部の機能を発揮しても良い。また、例えば、ＳＭ１２０とＣＭ１３０が、一体的に構成されていても良い。また、各チャネル制御部毎に１つのＬＵＮが割当てられても良いし、複数のチャネル制御部に１つのＬＵＮが割当てられても良い。また、記憶制御システム６００には、別の記憶制御システムが接続されていても良い。その場合、例えば、記憶制御システム６００が有するプライマリＬＵと、別の記憶制御システムが有するセカンダリＬＵとがペア状態にされて、記憶制御システム６００に接続されているホスト端末２００Ａ～２００Ｄが、記憶制御システム６００を介して別の記憶制御システム内のセカンダリＬＵにアクセスしても良い。また、例えば、記憶制御システム６００は、ファイルアクセス要求及びブロックアクセス要求の双方を受けるものであっても良いし、ファイルアクセス要求のみを受けるもの（例えばＮＡＳ）であっても良いし、ブロックアクセス要求のみを受けるものであっても良い。

【００３９】

図３は、本実施形態に係る記憶制御サブシステム１０２の機能を示すブロック図である。なお、以下の説明では、説明を分かり易くするために、ホスト端末２００Ａ～２００Ｄのうちホスト端末２００Ａを例に採り、且つ、チャネル制御部１１０Ａ～１１０Ｄのうちチャネル制御部１１０Ａを例に採る。

【００４０】

ホスト端末２００Ａと記憶制御サブシステム１０２との間には、１又は複数の論理的な通信パス（以下、「論理パス」と言う）２１Ａ～２１Ｄが形成されている。各論理パス２１Ａ～２１Ｄは、例えば、記憶制御サブシステム１０２が有するポート（ホスト端末１００Ａが接続されるポート）の番号と、ターゲットＩＤと、ＬＵＮとに基づいて形成されたものである。

【００４１】

１又は複数のＲＡＩＤグループ２上に用意される複数のＬＤＥＶには、例えば、通常ＬＤＥＶという属性を持ったＬＤＥＶ（以下、通常ＬＤＥＶ）６１ａと、プールＬＤＥＶという属性を持ったＬＤＥＶ（以下、プールＬＤＥＶ）６１ｂとがある。各ＬＤＥＶは、ＳＶＰ１６０のオペレータの指示により、通常ＬＤＥＶ６１ａからプールＬＤＥＶ６１ｂになることもできるし、逆に、プールＬＤＥＶ６１ｂから通常ＬＤＥＶ６１ａになることもできる。

【００４２】

通常ＬＤＥＶ６１ａは、ホスト端末１１０Ａがアクセス可能なＬＤＥＶである。換言すれば、例えば、論理パス２１Ａが有するＬＵＮがホスト端末１１０Ａから指定された場合、そのＬＵＮに対応付けられている２つの通常ＬＤＥＶ６１ａが、ＬＵ３１０Ａとしてホスト端末２００Ａに提供される。また、例えば、論理パス２１Ｂが有するＬＵＮがホスト

端末１１０Ａから指定された場合、そのＬＵＮに対応付けられている１つの通常ＬＤＥＶ６１ａが、ＬＵ３１０Ｂとしてホスト端末２００Ａに提供される。

【００４３】

プールＬＤＥＶ６１ｂは、ＬＤＥＶプール６８を構成するメンバであり、ホストがアクセス不可能なＬＤＥＶである。換言すれば、プールＬＤＥＶ６１ｂは、ホスト端末２００Ａから指定され得るＬＵＮに対応付けられておらず、ＬＵＮが指定されてもプールＬＤＥＶ６１ｂそれ自体はホスト端末２００Ａに提供されないようになっている。

【００４４】

ＬＤＥＶプール６８の上位には、ホスト端末２００Ａに対して提供される１又は複数の仮想ＬＵ３１０Ｃが存在する。仮想ＬＵ３１０Ｃは、ホスト端末２００Ａに提供されるＬＵであるが、他のＬＵ３１０Ａ、３１０Ｂと違い、RAIDグループ２上に物理的なデータ格納領域を持たない仮想的なＬＵである。具体的に言えば、仮想ＬＵ３１０Ｃは、仮想ＬＤＥＶ６１ｃから構成されるが、その仮想ＬＤＥＶ６１ｃは、通常ＬＤＥＶ６１ａやプールＬＤＥＶ６１ｂと異なり、RAIDグループ２上に物理的なデータ格納領域を持たない仮想的なＬＤＥＶである。仮想ＬＤＥＶ６１ｃが有する複数の仮想アドレスの各々は、ダイナミックマップテーブル（以下、「ＤＭＴ」と略記）６４を介して、動的に、プールＬＤＥＶ６８が有する複数の論理アドレスから選択された論理アドレスに対応付けられたり、その対応付けが解除されてその論理アドレスが開放されたりする。

【００４５】

図４は、仮想ＬＤＥＶ６１ｃ、ＬＤＥＶプール６８、及びＤＭＴ６４の構成例を示す。

【００４６】

仮想ＬＤＥＶ６１ｃは、例えば、一定サイズ（例えば６４キロバイト）を有する複数の仮想チャンク４１０ｃ、４１０ｃ、…から成っている。各仮想チャンク４１０ｃは、所定個数（例えば１２８個）の論理ブロック（例えば１個が５１２バイト）から構成されている。各仮想チャンク４１０ｃには、先頭論理ブロックアドレス（以下、「仮想先頭ＬＢＡ」と略記）が存在し、仮想先頭ＬＢＡから、どの仮想チャンク４１０ｃであるかを特定することができるようになっている。

【００４７】

各ＬＤＥＶプール６８は、１以上のＬＤＥＶ６１ｂの集合であり、例えば、一定サイズ（例えば６４キロバイト）を有する複数の論理チャンク４１０ｂ、４１０ｂ、…から成っている。各論理チャンク６１ｂは、所定個数（例えば１２８個）の論理ブロック（例えば１個が５１２バイト）から構成されている。各論理チャンク４１０ｂには、先頭論理ブロックアドレス（以下、「論理先頭ＬＢＡ」と略記）が存在し、論理先頭ＬＢＡから、どの論理チャンク４１０ｃであるかを特定することができるようになっている。なお、例えば、アクセスする論理先頭ＬＢＡが決まれば、対応するチャンクの番号が決まる。具体的には、例えば、論理チャンクサイズを6.4KB(＝１２８ブロック)とすると、論理チャンク番号ｎは、論理ブロック番号ｍ／１２８の整数部となる（例えば、論理ブロック番号が０～１２７の論理ブロックを有する論理チャンク番号はゼロであり、論理ブロック番号が１２８～２５５の論理ブロックを有する論理チャンクの番号は１である）。

【００４８】

ＤＭＴ６４は、外部端末（例えばＳＶＰ１６０）から記憶制御サブシステム１０２内の記憶領域（例えばＳＭ１２０）に登録される。ＤＭＴ６４は、各仮想ＬＤＥＶ６１ｃ毎に用意され、その仮想ＬＤＥＶ６１ｃの各仮想チャンク４１０ｃと、１以上のＬＤＥＶプール６８の各々における各論理チャンク４１０ｂとを対応付けるためのテーブルである。具体的には、例えば、ＤＭＴ６４には、各仮想先頭ＬＢＡ毎に、その仮想先頭ＬＢＡに対応付けられたＬＤＥＶプール６８の識別情報（例えばＬＤＥＶプール番号）と、そのＬＤＥＶプール６８における論理先頭ＬＢＡとが記述される。このＤＭＴ６４の記述内容は、所定のタイミング、例えば、チャネル制御部１１０ＡがＩ／Ｏ要求を受けた場合に、チャネル制御部１１０Ａによって更新される。チャネル制御部１１０Ａは、プールマネジメントテーブル（以下、「ＰＭＴ」と略記）６３を参照して、ＤＭＴ６４の記述内容の更新を更

新する。

【００４９】

図５は、ＰＭＴ６３の構成例を示す。

【００５０】

ＰＭＴ６３は、ＬＤＥＶプール６８毎に存在するものであり、外部端末（例えばＳＶＰ１６０）から記憶制御サブシステム１０２内の記憶領域（例えばＳＭ１２０）に登録される。ＬＤＥＶプール６８の各論理チャンク４０１ｂ毎に、エントリ番号が割り当てられる。各ＰＭＴ６３には、それに対応したＬＤＥＶプール６８における初めの空きエントリの番号（換言すれば、論理先頭ＬＢＡが最も若いチャンク番号）を筆頭に、キュー形式で、他の空きエントリの番号が書かれる。

【００５１】

以下、仮想ＬＤＥＶ６１ｃとＬＤＥＶプール６８とを動的に対応付けるチャネル制御部１１０Ａの処理流れについて説明する。

【００５２】

ホスト端末２００Ａは、通常のＬＵ３１０Ａにアクセスする場合と同様の方法で、仮想ＬＵ３１０Ｃに対するＩ／Ｏ要求を出力する。チャネル制御部１１０Ａは、例えば、ホスト端末１１０Ａからそのの要求を受けた場合、そのＩ／Ｏ要求から割り出された仮想先頭ＬＢＡと、ＬＤＥＶプール６８の論理先頭ＬＢＡとの対応付けを行う。具体的には、例えば、チャネル制御部１１０Ａは、Ｉ／Ｏ要求を受けた仮想ＬＵ３１０Ｃが有する仮想ＬＤＥＶ６１ｃに対応したＤＭＴ６４を参照し、上記割り出された仮想先頭ＬＢＡに対応したＬＤＥＶプール番号を取得する。次に、チャネル制御部１１０Ａは、取得したＬＤＥＶプール番号に対応したＰＭＴ６３から、そのＰＭＴ６３に書かれている最初の空きエントリ番号に対応した論理先頭ＬＢＡを取得し、その論理先頭ＬＢＡを、上記仮想先頭ＬＢＡに対応した場所（ＤＭＴ６４上の場所）に登録する。それにより、上記仮想先頭ＬＢＡに論理先頭ＬＢＡが対応付けられ、その論理先頭ＬＢＡを持つ論理チャンクからデータが読み出されたり、その論理チャンクにデータが書き込まれたりすることができるようになる。

【００５３】

また、チャネル制御部１１０Ａは、例えば、上記Ｉ／Ｏ要求に基づく処理の終了後に、ＤＭＴ６４に書かれた論理先頭ＬＢＡを開放することができる。具体的には、例えば、チャネル制御部１１０Ａは、DMT６４に登録した上記論理先頭ＬＢＡをＤＭＴ６４から削除し、ＰＭＴ６３に、その論理先頭ＬＢＡ及びそれに対応するエントリ番号をＰＭＴ６４に格納する。

【００５４】

以上のような流れで、仮想ＬＵ３１０Ｃ（仮想ＬＤＥＶ６１ｃ）とＬＤＥＶプール６８との動的な対応付けやそれの解除が行われる。このように、どのホスト端末からアクセスされるかが予め割当てられていないプールＬＤＥＶ６１ｂ、６１ｂ、…から成るＬＤＥＶプール６８を用意し、そのＬＤＥＶプール６８における記憶領域をフレキシブルに仮想ＬＵ３１０Ｃ（仮想ＬＤＥＶ６１ｃ）に対応付けることにより、空いている記憶領域を有効に活用することができる。

【００５５】

図６は、この実施形態に係るＬＵ－ＬＤＥＶ管理テーブル１６２ｂの構成例を示す。

【００５６】

ＬＵ－ＬＤＥＶ管理テーブル１６２ｂは、記憶制御サブシステム１０２内の記憶領域（例えばＳＭ１２０）に登録されるものである。ＬＵ－ＬＤＥＶ管理テーブル１６２ｂには、例えば、記憶制御システム６００が備える各ポート（ホスト端末が接続されるポート）毎に、ポート番号と、そのポート番号に属する１以上のターゲットＩＤと、各ターゲットＩＤに属する１以上のＬＵＮと、各ＬＵＮに属する１以上のＬＤＥＶに関するＬＤＥＶ情報が記録されている。各ＬＤＥＶのＬＤＥＶ情報としては、例えば、ＬＤＥＶ番号、記憶容量、ＲＡＩＤレベル、上述したアドレス管理情報、ＬＤＥＶ属性、ＤＭＴ－ＩＤ、ＬＤ

ＥＶプール番号；状態及び記憶容量通知済みホストＩＤがある。

【００５７】

ＬＤＥＶ属性としては、例えば、上述したプールＬＤＥＶ、通常ＬＤＥＶ及び仮想ＬＤＥＶというものがある。

【００５８】

「プールＬＤＥＶ」というＬＤＥＶ属性を持つＬＤＥＶ（前述したプールＬＤＥＶ６１ｂ）は、ホスト端末にアクセスされないようになっているため、プールＬＤＥＶ６１ｂには、論理パス情報（例えばＬＵＮ）は対応付けられていない。また、プールＬＤＥＶ６１ｂには、どのＬＤＥＶプール６８のメンバであるかが識別されるようにＬＤＥＶプール番号が対応付けられている。

【００５９】

「仮想ＬＤＥＶ」というＬＤＥＶ属性を有するＬＤＥＶ（前述した仮想ＬＤＥＶ６１ｃ）には、そのＬＤＥＶ６１ｃに対応するＤＭＴ６４のＩＤ（例えばゼロ）が対応付けられる。このＤＭＴ－ＩＤを参照することにより、どの仮想ＬＤＥＶ６１ｃ（仮想ＬＵ３１０Ｃ）にアクセスされた場合にはどのＤＭＴを参照すれば良いかがわかるようになっている。また、仮想ＬＤＥＶ６１ｃには、どのＬＤＥＶプール６８のメンバであるかが識別されるようにＬＤＥＶプール番号が対応付けられている。

【００６０】

「通常ＬＤＥＶ」というＬＤＥＶ属性を有するＬＤＥＶ（前述した通常ＬＤＥＶ６１ａ）には、ＤＭＴ－ＩＤ及びＬＤＥＶプール番号は対応付けられていない。

【００６１】

ＬＤＥＶの状態としては、例えば、アクセス可能な状態であることを表す「Ready」（例えば実装状態）と、アクセス不可能な状態であることを表す「Not Ready」（例えば未実装状態）とがある。

【００６２】

各ＬＤＥＶの記憶容量通知済みホストＩＤとは、ＬＤＥＶが属するＬＵの記憶容量の通知を受けたホスト端末のＩＤ（例えば、ＭＡＣアドレス又はWWN（World Wide Name）である。これを参照することにより、どのＬＵの記憶容量がどのホスト端末に通知されたかが分かるようになっている。

【００６３】

図７は、チャネル制御部１１０Ａの構成例を示す。

【００６４】

チャネル制御部１１０Ａは、ハードウェア要素が一体的にユニット化されたボードで構成される。チャネル制御部１１０Ａは、ＣＨＮ（Channel adapter Nas）又はＮＡＳブレードと呼ばれることがある。チャネル制御部１１０Ａは、図示しないボード接続用コネクタを備えており、そのボード接続用コネクタが記憶制御装置１００の所定コネクタに嵌合することにより、チャネル制御部１１０Ａは記憶制御装置１００と電気的に接続される。チャネル制御部１１０Ａは、例えば、ホストインターフェース部（以下、「ホストＩ／Ｆ」と記載）７１１、ＳＶＰインターフェース部（以下、「ＳＶＰＩ／Ｆ」と記載）５１、ブリッジＬＳＩ（Large-Scale Integration）７８１、メモリコントローラ７４１、ＮＡＳプロセッサ１１２、ホスト側メモリ１１３、１又は複数の入出力制御部７７１及びデータ転送ＬＳＩ７８２を備える。

【００６５】

ホストＩ／Ｆ７１１は、ホスト端末２００Ａ及び２００Ｂとの間で通信を行うための通信インタフェースであり、例えば、ＴＣＰ／ＩＰプロトコルに従ってホスト端末２００Ａ及び２００Ｂから送信されたファイルアクセス要求を受信する。

【００６６】

ＳＶＰＩ／Ｆ５１は、内部ＬＡＮ１５０等の通信ネットワークを介してＳＶＰ１６０に接続され、且つ、後述のＣＨＰ１１９に接続される。ＳＶＰＩ／Ｆ５１は、ＳＶＰ１６０とＣＨＰ１１９との間の通信を制御するための通信インタフェース（例えばＬＡＮコント

ローラ）である。

【００６７】

ブリッジＬＳＩ７８１は、例えば、ホストＩ／Ｆ７１１、メモリコントローラ７４１及びデータ転送ＬＳＩ７８２の間の通信を可能にするためのＬＳＩである。

【００６８】

メモリコントローラ５０５は、ブリッジＬＳＩ７８２、ＮＡＳプロセッサ１１２及びホスト側メモリ１１３の間の通信を可能にするためのＬＳＩである。

【００６９】

ＮＡＳプロセッサ１１２は、例えばＣＰＵであり、ＮＦＳ（Network File System）を用いて、記憶制御サブシステム６００をＮＡＳとして機能させるための制御を行う。例えば、ＮＡＳプロセッサ１１２は、ＮＦＳ等により、ファイル名と論理ブロックアドレスとの対応関係を把握しており、チャネル制御部１１０Ａが受けたファイルアクセス要求に含まれているファイル名及び上記対応関係を基に、そのファイルアクセス要求をブロックアクセス要求に変換して、ＣＨＰ１２１に出力する。

【００７０】

ホスト側メモリ１１３には、例えば、様々なプログラムやデータが記憶される。具体的には、例えば、ホスト側メモリ１１３には、後述するファイルメタデータテーブル、ロックテーブル、及びＮＡＳマネージャ等のデータやプログラムが記憶される。

【００７１】

各入出力制御部７７１は、ディスク制御部１４０Ａ～１４０Ｄ、ＣＭ１３０、ＳＭ１２０及びＳＶＰ１６０との間でデータやコマンドの授受を行うもの（例えばマイクロプロセッサユニット）である。各入出力制御部７７１は、ＣＨＰ（チャネルプロセッサ）１２１や、ＣＨＰメモリ１１９を備える。

【００７２】

ＣＨＰ１２１は、例えば、マイクロプロセッサであり、チャネル制御部１１０Ａ全体の制御を司ると共に、ディスク制御部１４０Ａ～１４０Ｄや、ホスト端末２００Ａ及び２００Ｂや、ＳＶＰ１６０との間の通信を制御する。ＣＨＰ１２１は、ＣＨＰメモリ１２１（又はホスト側メモリ１１３）に格納された各種コンピュータプログラムを実行することにより、チャネル制御部１１０Ａとしての機能を発揮する。

【００７３】

ＣＨＰメモリ１１９は、揮発性又は不揮発性のメモリ（例えばＮＶＲＡＭ（Non Volatile RAM））であり、例えば、ＣＨＰ１２１の制御を司るコンピュータプログラムを格納する。ＣＨＰメモリ１１９に記憶されるプログラムの内容は、例えば、ＳＶＰ１６０や、後述するＮＡＳマネージャ８０６からの指示により書き込みや書き換えを行うことができる。

【００７４】

データ転送ＬＳＩ５０１は、データ転送の制御を行うためのＬＳＩである。具体的には、例えば、データ転送ＬＳＩ５０１は、ＣＭデータ転送回路７１０及びＳＭデータ転送回路７４０を備える。ＣＭデータ転送回路７１０は、ＣＭ１３０に接続され、ユーザデータの入出力を行う。ＳＭデータ転送回路７４０は、ＳＭ１２０に接続され、制御情報の入出力を行う。

【００７５】

上述したホスト側メモリ１１３（又はＣＨＰメモリ１２１）には、例えば、リードキャパシティ処理プログラム７１５及びＬＵ定義処理プログラム７１６が格納される。

【００７６】

リードキャパシティ処理プログラム７１５は、ホスト端末２００Ａからリードキャパシティコマンドを受信した場合に、そのリードキャパシティコマンドが受けたポート番号、ターゲットＩＤ及びＬＵＮに対応したＬＵの記憶容量をＬＵ－ＬＤＥＶ管理テーブル１６２ｂから取得し、その記憶容量を、リードキャパシティコマンド発行元のホスト端末２００Ａに通知する。

【００７７】
ＬＵ定義処理プログラム７１６は、ＳＶＰ１６０からのアクセスに応答して、ＬＵ設定画面をＳＶＰ１６０に提供し、そのＬＵ設定画面に入力された情報に基づいて、ＬＵ－ＬＤＥＶ管理テーブル１６２ｂを更新する。
【００７８】
なお、リードキャパシティ処理プログラム７１５及びＬＵ定義処理プログラム７１６の少なくとも一方は、必ずしもチャネル制御部１１０Ａに格納されている必要は無く、例えばディスク制御部１４０Ａ～１４０Ｄに搭載されているメモリ等別の場所に格納されていても良い。
【００７９】
図８は、ＬＵが設定される場合に行なわれる記憶制御サブシステム１０２での処理流れを示す。
【００８０】
この図に示す処理は、例えば、チャネル制御部１１０Ａが実行する。
【００８１】
チャネル制御部１１０Ａは、ＳＶＰ１６０（又はそれに接続された外部端末）から、ＬＤＥＶプールの作成依頼を受ける。チャネル制御部１１０Ａは、その依頼に応答して、ＬＤＥＶプール編集画面５５２をＳＶＰ１６０に提供する。この画面５５２には、例えば、作成対象のＬＤＥＶプール番号の入力欄と、そのＬＤＥＶプールに対して追加、除去又は編集するためのＬＤＥＶ番号の入力欄とがある。また、例えば、その画面５５２には、ＬＤＥＶ番号の入力に入力されたＬＤＥＶ番号を追加、除去又は編集のいずれを行うかの選択を受け付けるツールが設けられている。
【００８２】
チャネル制御部１１０Ａは、そのＬＤＥＶプール編集画面５５２に入力された情報に従って、ＬＤＥＶプールの作成（又は編集）を行う（換言すれば、ＬＵ－ＬＤＥＶ管理テーブル１６２ｂを更新する）（ステップＳ３００）。
【００８３】
その後、チャネル制御部１１０Ａは、仮想ＬＵと通常ＬＵのいずれを定義するかの選択をＳＶＰ１６０から受付け、仮想ＬＵが選択された場合、仮想ＬＵ設定画面５５１をＳＶＰ１６０に提供する。その画面５５１には、その仮想ＬＵに対応付けるポート番号、ターゲットＩＤ、ＬＵＮ、ＬＤＥＶ番号及び仮想記憶容量の入力欄がある。
【００８４】
チャネル制御部１１０Ａは、その仮想ＬＵ設定画面５５１に入力された情報に従って、仮想ＬＵを定義する（換言すれば、「仮想ＬＤＥＶ」というＬＤＥＶ属性を持った新たな仮想ＬＤＥＶに関する情報をＬＵ－ＬＤＥＶ管理テーブル１６２ｂに登録する）（Ｓ３１０）。
【００８５】
また、チャネル制御部１１０Ａは、Ｓ３１０で定義した仮想ＬＵがどのＬＤＥＶプールを使用するかの指定を受け（例えば、そのＬＤＥＶプールの番号の入力を受け）、指定されたＬＤＥＶプールを仮想ＬＵ（仮想ＬＤＥＶ）に対応付ける（換言すれば、入力されたＬＤＥＶプール番号をＬＵ－ＬＤＥＶ管理テーブル１６２ｂに登録する）（Ｓ３２０）。
【００８６】
以上の流れにより、仮想ＬＵ（仮想ＬＤＥＶ）の定義付けが完了する。なお、ここで入力された仮想記憶容量は、ホスト端末２００Ａからのリードキャパシティコマンドに応答して、そのホスト端末２００Ａに通知される。
【００８７】
図９は、ホスト端末２００Ａの処理フローの一例を示す。
【００８８】
ホスト端末２００Ａは、例えば、再起動処理を実行する場合、又は、ホスト端末２００Ａに接続されている全ての記憶制御システム６００を調査する再スキャン処理を実行する

場合、以下の処理フローを実行する。

【００８９】

例えば、ホスト端末２００Ａは、第1のコマンド（例えばInquiryコマンド）を記憶制御システム６００に発行して、その記憶制御システム６００に関する制御システム情報（例えば、その記憶制御システム６００のベンダ名やモデル名）をその記憶制御システム６００から受信する（Ｓ２０１）。

【００９０】

次に、ホスト端末２００Ａは、第２のコマンド（例えばデバイスディスカバリーコマンド又はリードキャパシティコマンド）を記憶制御システム６００に発行して、その記憶制御システム６００がホスト端末２００Ａに提供する全てのＬＵに関するＬＵ情報を、その記憶制御システム６００から受信する（Ｓ２０２）。受信するＬＵ情報には、ホスト端末２００Ａがアクセス可能な各ＬＵ毎に、例えば、少なくとも記憶容量情報が含まれており、その他、そのＬＵのパス（例えば、ポート番号、ターゲットＩＤ及びＬＵＮ）や、そのＬＵに属するＬＤＥＶを持ったディスク型記憶装置のベンダ等も含まれていても良い。

【００９１】

ホスト端末２００Ａは、Ｓ２０１で受信した制御システム情報や、Ｓ２０２で受信したＬＵ情報を、ホスト端末２００Ａの記憶装置（例えばハードディスクドライブ）に格納する（Ｓ２０３）。例えば、ホスト端末２００Ａは、Ｓ２０２で受信したＬＵ情報に基づいて、各記憶制御システム６００毎に、図１０に例示するような記憶制御システム管理テーブル７９１を作成して記憶する。記憶制御システム管理テーブル７９１には、各ＬＵ毎に、ポート番号、ターゲットＩＤ、ＬＵＮ、ベンダ、モデル及び記憶容量等が登録される。なお、ＬＵが仮想ＬＵの場合、Ｓ２０２で受信する記憶容量は、仮想ＬＵ設定画面５５１で入力された値（すなわち、オペレータ任意の値）である。

【００９２】

ホスト端末２００Ａは、全ての記憶制御システム６００に対して、Ｓ２０１～Ｓ２０３処理を実行する（Ｓ３０４でＮ）。

【００９３】

図１１は、チャネル制御部１１０Ａが、ホスト端末２００から第２コマンドとしてリードキャパシティコマンドを受けた場合に行なう処理フローの一例を示す。

【００９４】

チャネル制御部１１０Ａは、ホスト端末２００Ａからリードキャパシティコマンドを受けた場合、そのリードキャパシティコマンドを受けたポート番号、ターゲットＩＤ及びＬＵＮに対応した記憶容量をＬＵ－ＬＤＥＶ管理テーブル１６２ｂから取得し（Ｓ２１１）、その記憶容量をホスト端末２００Ａに通知する（Ｓ２１２）。また、チャネル制御部１１０Ａは、通知先ホスト端末２００ＡのホストＩＤを、ＬＵ－ＬＤＥＶ管理テーブル１６２ｂに書き込む。

【００９５】

以上の説明では、説明を分かり易くするために、チャネル制御部１１０Ａ及びホスト端末２００Ａを例に採り説明したが、上記説明は、その他のチャネル制御部１１０Ｂ～１１０Ｄ、ディスク制御部１４０Ａ～１４０Ｄ、及びその他のホスト端末１２０Ｂ～１２０Ｄにも当てはまるものである。

【００９６】

ところで、本実施形態では、更に、例えば以下の特徴的な処理が行われる。以下、上述の説明と重複するかもしれないが、幾つかの特徴的な処理について説明する。

【００９７】

（１）第1の特徴的な処理

図１２は、第１の特徴的な処理に関わる処理流れを示す。

【００９８】

記憶制御サブシステム１０２の記憶制御装置１００（例えば、チャネル制御部又はディスク制御部）は、仮想ＬＵ３１０Ｃの仮想記憶容量値を所定記憶領域（例えばＳＭ１２０

）に登録した場合に、仮想ＬＵ３１０Ｃを実装状態にする（換言すれば、仮想ＬＵ３１０Ｃをホスト端末に対してアクセス可能にする）（Ｓ１）。

【００９９】

その後、記憶制御装置１００は、ＳＶＰ１６０（又はその他の外部端末）から、仮想記憶容量値の変更依頼を受けた場合（Ｓ２）、Ｓ１で登録した仮想記憶容量値をホスト端末に通知済みか否かを判別する（Ｓ３）。具体的には、例えば、記憶制御装置１００は、ＬＵ－ＬＤＥＶ管理テーブル１６２ｂを参照し、変更依頼対象の記憶容量値に対応した通知済みホストＩＤが登録されているか否かを判別する。

【０１００】

Ｓ３の結果、肯定的な判別結果が得られた場合（Ｓ３でＹ）、記憶制御装置１００は、ＳＶＰ１６０からの記憶容量値変更依頼を拒否する（Ｓ４）。すなわち、記憶制御装置１００は、一度でも過去にホスト端末に仮想記憶容量値を通知していれば、仮想記憶容量値の変更を禁止する。

【０１０１】

一方、Ｓ３の結果、否定的な判別結果が得られた場合（Ｓ３でＮ）、記憶制御装置１００は、ＳＶＰ１６０からの記憶容量値変更依頼を受けると共に、ＳＶＰ１６０から変更後の仮想記憶容量値の入力を受け、それを、所定記憶領域における旧い仮想記憶容量値に上書きする（Ｓ５）。

【０１０２】

以上の処理により、一度でも過去にホスト端末に仮想記憶容量値が通知されて、そのホスト端末でその仮想記憶容量値が記憶されている可能性があれば、同一の仮想ＬＵについて別の仮想記憶容量値がそのホスト端末に通知されて、そのホスト端末を混乱させてしまわないようなことがないように制御される。

【０１０３】

（２）第２の特徴的な処理

記憶制御装置１００は、オペレータに任意に設定された仮想記憶容量値を所定記憶領域に登録し、ホスト端末から第２コマンドを受信したら、そのコマンドに応じて、上記オペレータに設定された任意値をホスト端末に通知する。

【０１０４】

（３）第３の特徴的な処理

図１３は、第３の特徴的な処理に関わる構成及び処理流れを示す。

【０１０５】

記憶制御装置１００は、２つのＬＵから成るＬＵペアを形成し、一方のＬＵをプライマリＬＵ３１０Ｐ、他方のＬＵをセカンダリＬＵ３１０Ｓとして、プライマリＬＵ３１０Ｐ内のデータをセカンダリＬＵ３１０Ｓにコピーするスナップショットを行うようになっている。

【０１０６】

具体的には、例えば、記憶制御装置１００は、外部端末（例えばＳＶＰ１６０）からの要求に手動で、又は、ライト要求を示すＩ／Ｏ要求を受領した場合（Ｓ１１）に自動で、通常ＬＵ３１０Ａをプライマリ ＬＵ３１０Ｐとし、仮想ＬＵ３１０Ｃをセカンダリ ＬＵ３１０ＳとしたＬＵペアを形成する。

【０１０７】

その後、記憶制御装置１００は、受領したライト要求から、通常ＬＵ３１０Ａにおける書込み先チャンク（例えば先頭ＬＢＡ）４１０Ｐを特定する（Ｓ１２）。そして、記憶制御装置１００は、特定されたチャンク４１０Ｐ内の一部のデータ（斜線で示した部分）しか更新されなくても、そのチャンク４１０Ｐ内の全てのデータを読み出し（Ｓ１３）、そのデータを、仮想ＬＵ３１０Ｃに書き込む（Ｓ１４）。具体的には、記憶制御装置１００は、前述したように、仮想ＬＵ３１０Ｃ内の仮想チャンク４１０Ｓと、ＬＤＥＶプール６８内の論理チャンクとを動的に対応付けて、そのデータを、仮想チャンク４１０Ｓに対応付けられたＬＤＥＶチャンクに書き込む。

【０１０８】
このようにして、スナップショットを作成した後、記憶制御装置１００は、上記特定されたチャンク４１０Ｐに、Ｓ１１で受領したライト要求に含まれているデータを書き込む（Ｓ１５）。
【０１０９】
なお、上記流れにおいて、記憶制御装置１００は、ＬＵペアを形成しようとする場合、ＬＵ－ＬＤＥＶ管理テーブル１６２ｂを参照し、プライマリＬＵ３１０Ｐとセカンダリ３１０Ｓとの記憶容量値が一致した場合にのみＬＵペアを形成し、それらが一致しない場合、ＬＵペアの形成を拒否しても良い。
【０１１０】
プライマリＬＵ３１０Ｐとセカンダリ３１０Ｓとの記憶容量値が同一でないとＬＵペアが形成されないようになっている場合、プライマリＬＵ３１０Ｐの記憶容量が大容量であると、セコンダリＬＵ３１０Ｓの記憶容量も大容量である必要がある。しかし、スナップショット作成では、プライマリＬＵ３１０Ｐからセコンダリ３１０Ｓにコピーするデータサイズが、プライマリＬＵ３１０Ｐの容量よりも小さい場合がある。それにもかかわらず、大容量の通常ＬＵ３１０Ａをセコンダリ３１０Ｓとしてしまうと、大量の無駄な空き記憶領域が生じてしまうことが考えられる。
【０１１１】
しかし、以上のように、スナップショットの作成において、セカンダリＬＵ３１０Ｓを仮想ＬＵ３１０ＣとしてＬＵペアを形成することにより、その仮想ＬＵ３１０Ｃに対応付けられ得るＬＤＥＶプール６８の空き領域は動的に別の仮想ＬＵ３１０Ｃに対応付けられるようになっているので、無駄に空き記憶領域が生じてしまうことを防ぐことができる。
【０１１２】
（４）第４の特徴的な処理
図１４は、第４の特徴的な処理に関わる処理流れを示す。
【０１１３】
記憶制御装置１００は、例えば、仮想記憶容量値を持たない仮想ＬＵ３１０Ｃを準備する。
【０１１４】
その後、記憶制御装置１００は、ＬＵペア形成コマンドを外部端末（例えばＳＶＰ１６０）から受信した場合（Ｓ３１）、そのＬＵペア形成コマンドで指定されたセカンダリＬＵ３１０Ｓである仮想ＬＵ３１０Ｃの仮想記憶容量値が所定記憶領域に登録されているか否か（例えば、ＬＵ－ＬＤＥＶ管理テーブル１６２ｂに記憶容量値の登録があるか否か）を判別する（Ｓ３２）。
【０１１５】
Ｓ３２の結果、肯定的な判別結果が得られれば（Ｓ３２でＹ）、記憶制御装置１００は、Ｓ３１で受信したＬＵペア形成コマンドに従って、ＬＵペアを形成する（Ｓ３４）。
【０１１６】
一方、Ｓ３２の結果、否定的な判別結果が得られれば（Ｓ３２でＮ）、記憶制御装置１００は、上記ＬＵペア形成コマンドで指定されたプライマリＬＵ３１０Ｐ（例えば通常ＬＵ３１０Ａ又は仮想ＬＵ３１０Ｃ）の記憶容量値を仮想記憶容量値として所定記憶領域に登録する（Ｓ３３）。その後、記憶制御装置１００は、Ｓ３４の処理を実行する。
【０１１７】
（５）第５の特徴的な処理
図１５は、第５の特徴的な処理に関わる処理流れを示す。
【０１１８】
記憶制御装置１００は、例えば、仮想記憶容量値を持たない仮想ＬＵ３１０Ｃを準備する。また、記憶制御装置１００は、各ＬＵペアに関するＬＵペア情報（例えばＬＵペアを構成する各ＬＵのＬＵＮ）を所定の記憶領域（例えばＳＭ１２０）で管理している。
【０１１９】

記憶制御装置１００は、そのような仮想ＬＵ３１０Ｃに対するＩ／Ｏ要求を受信した場合（Ｓ４１）、上記ＬＵペア情報を参照して、その仮想ＬＵのペア相手が存在するか否かを判別する（Ｓ４２）。

【０１２０】

Ｓ４２の判別の結果、否定的な判別結果が得られた場合（Ｓ４２でＮ）、記憶制御装置１００は、ＬＵが未実装である旨を、Ｉ／Ｏ要求送信元のホスト端末に通知し（Ｓ４７）、処理を終える。なお、その後、記憶制御装置１００は、その仮想ＬＵの相手ＬＵが設定された後、所定のタイミングで（例えば、ペア相手ＬＵが設定された時点で）、その相手ＬＵの記憶容量値を仮想ＬＵの仮想記憶容量値として所定記憶領域に登録する。また、記憶制御装置１００は、ＬＵが未実装である旨の通知先となったホスト端末のホストＩＤを記憶しておき、仮想記憶容量値を登録した時点で、その仮想記憶容量値を、そのホストＩＤに対応したホスト端末に通知しても良い。

【０１２１】

一方、Ｓ４２の判別の結果、肯定的な判別結果が得られた場合（Ｓ４２でＹ）、記憶制御装置１００は、仮想記憶容量値が所定記憶領域に登録されているか否かを判別する（Ｓ４３）。

【０１２２】

このＳ４３の判別の結果、肯定的な結果が得られた場合（Ｓ４３でＹ）、記憶制御装置１００は、Ｓ４１で受信したＩ／Ｏ要求に従う処理を実行する（Ｓ４５）。

【０１２３】

一方、Ｓ４３の判別の結果、否定的な結果が得られた場合（Ｓ４３でＮ）、記憶制御装置１００は、仮想ＬＵの相手ＬＵの記憶容量値を仮想記憶容量値として所定記憶領域に登録すると共に、その仮想記憶容量値を、Ｉ／Ｏ要求送信元のホスト端末に通知する（Ｓ４４）。その後、記憶制御装置１００は、Ｓ４５の処理を実行する。

【０１２４】

（６）第６の特徴的な処理

図１６は、第６の特徴的な処理に関わる処理流れを示す。

【０１２５】

ホスト端末は、所定の処理、例えば、再起動処理を実行する場合、又は、そのホスト端末に接続されている全ての記憶制御システム６００を調査する再スキャン処理を実行する場合、過去に記憶した情報（例えば、記憶制御システム管理テーブルそれ自体）を消去するようになっている。

【０１２６】

記憶制御装置１００は、例えば、外部端末（例えばＳＶＰ１６０）から、仮想記憶容量値の更新依頼を受けた場合（Ｓ５１）、その仮想記憶容量値を持つ仮想ＬＵ３１０Ｃにアクセス可能な全てのホスト端末との接続を切断する（換言すれば、その仮想ＬＵ３１０Ｃを強制的にオフライン状態にする）（Ｓ５２）。

【０１２７】

そして、記憶制御装置１００は、外部端末から新たな仮想記憶容量値の入力を受け、旧い仮想記憶容量値にその新たな仮想記憶容量値を上書きする（Ｓ５３）。

【０１２８】

その後、記憶制御装置１００は、接続が切断されたホスト端末に対して再起動処理又は再スキャン処理命令を出力する（Ｓ５４）。

【０１２９】

再起動処理又は再スキャン処理命令を受けたホスト端末は、その命令に応答して再起動処理又は再スキャン処理を実行する（Ｓ５５）。その際、ホスト端末は、過去に記憶した情報（例えば、記憶制御システム管理テーブルそれ自体）を消去する。

【０１３０】

Ｓ５５の処理の完了後、ホスト端末は、処理終了通知を記憶制御装置１００に送信する（Ｓ５６）。なお、この処理終了通知に加えて又は代えて、上述した第１コマンド（例え

ばInquiryコマンド）や、第２コマンド（例えばデバイスディスカバリーコマンド又はリードキャパシティコマンド）が送信されても良い。

【０１３１】

記憶制御装置１００は、ホスト端末から処理終了通知を受けたら、Ｓ５３で上書きした新たな仮想記憶容量値を通知する（Ｓ５７）。

【０１３２】

ホスト端末は、記憶制御装置１００から通知された新たな仮想記憶容量値を記憶する（Ｓ５８）。

【０１３３】

以上の処理流れにおいて、記憶制御装置１００は、Ｓ５３とＳ５４の処理の間に、旧い仮想記憶容量値をホスト端末に通知済みか否かを判断する処理を行っても良い。その場合、肯定的な判断結果が得られた場合には、記憶制御装置１００は、Ｓ５４の処理を実行し、否定的な判断結果が得られた場合には、Ｓ５４の処理を実行せずにＳ５７の処理を行っても良い。

【０１３４】

以上、本発明の実施形態を説明したが、これは本発明の説明のための例示であって、本発明の範囲をこの実施形態にのみ限定する趣旨ではない。本発明は、他の種々の形態でも実施することが可能である。

【０１３５】

例えば、本発明に従う記憶制御サブシステムの第１実施態様では、前記記憶制御サブシステムの保守のための処理を行う保守用端末が前記記憶制御部に接続されている場合において、前記記憶制御部は、前記保守用端末又は前記保守用端末に接続された外部端末から、新たな前記仮想記憶ユニットを準備することのユニット準備要求を受け、前記ユニット準備要求に応答して、少なくとも前記仮想記憶容量値の記入欄を持ったグラフィカルユーザインターフェースを前記保守用端末又は前記外部端末に提供し、前記記入欄に入力された仮想記憶容量値を、前記設定された仮想記憶容量値として前記メモリに記憶させる。

【０１３６】

例えば、本発明に従う記憶制御サブシステムの第２実施態様では、前記記憶制御サブシステムが、２つの記憶ユニットから成るユニットペアを形成し、一方の記憶ユニットをプライマリ記憶ユニット、他方の記憶ユニットをセカンダリ記憶ユニットとして、プライマリ記憶ユニット内のデータをセカンダリ記憶ユニットにコピーするスナップショットを行うようになっている場合、前記物理記憶デバイスには複数の前記論理記憶デバイスが設けられており、前記複数の論理記憶デバイスには、前記仮想記憶領域に対応付けられ得る論理記憶領域を持った２以上の第１論理記憶デバイスと、前記仮想記憶領域に対応付けられることのない論理記憶領域を持った１以上の第２論理記憶デバイスとが含まれており、前記１以上の第２論理記憶デバイスが、前記ホスト端末に接続される１つのリアル記憶ユニットを構成し、前記記憶制御部は、前記リアル記憶ユニットを前記プライマリ記憶ユニットとし、前記仮想記憶ユニットをセカンダリ記憶ユニットとしたユニットペアを形成して前記スナップショットを行う。

【０１３７】

例えば、本発明に従う記憶制御サブシステムの第３実施態様では、前記第２実施態様において、前記記憶制御部は、前記仮想記憶容量値を前記ホスト端末に通知していない場合に、前記仮想記憶ユニットと前記リアル記憶ユニットの前記ユニットペアを形成するならば、前記リアル記憶ユニットの記憶容量値と同じ値を前記仮想記憶ユニットの記憶容量値として前記ホスト端末に通知する。

【０１３８】

例えば、本発明に従う記憶制御サブシステムの第４実施態様では、前記第３実施態様において、前記記憶制御部は、前記仮想記憶ユニットの相手となる前記リアル記憶ユニットが見つからない場合に、前記ホスト端末から前記仮想記憶ユニットに対してリード要求又はライト要求を受けたならば、前記仮想記憶ユニットが未実装状態であると前記ホスト端

末に通知し、その後、前記相手となるリアル記憶ユニットが見つかったならば、前記リアル記憶ユニットの記憶容量値と同じ値を前記仮想記憶ユニットの記憶容量値として前記ホスト端末に通知する。

【０１３９】

例えば、本発明に従う記憶制御サブシステムの第５実施態様では、前記記憶制御部は、前記ホスト端末からリードキャパシティコマンドを受信した場合に、前記メモリに記憶されている仮想記憶容量値を前記ホスト端末に通知する。

【０１４０】

例えば、本発明に従う記憶制御サブシステムの第６実施態様では、前記記憶制御サブシステムの保守のための処理を行う保守用端末が前記記憶制御部に接続されており、且つ、前記ホスト端末が行う起動処理又は再スキャン処理で、前記ホスト端末が、前記記憶した仮想記憶容量値を消去するようになっている場合において、前記記憶制御部は、前記仮想記憶容量値を前記ホスト端末に通知した後に、前記保守用端末又は前記保守用端末に接続された外部端末から、前記通知した仮想記憶容量値の更新要求を受けたならば、前記ホスト端末と前記仮想記憶ユニットとが接続されていない間に、前記保守用端末又は前記外部端末から新たな仮想記憶容量値を受けて前記メモリに記憶させ、且つ、前記ホスト端末に前記起動処理又は前記再スキャン処理を行わせることで、前記ホストに記憶されている旧い前記仮想記憶容量値が消去されるようにした後、前記メモリに記憶させた新たな仮想記憶容量値を前記ホスト端末に通知する。

【図面の簡単な説明】

【０１４１】

【図１】本発明の一実施形態に係る記憶制御システムの外観の概略を示す。

【図２】本発明の一実施形態に係る記憶システムの全体構成例を示す。

【図３】本実施形態に係る記憶制御サブシステムの機能を示すブロック図。

【図４】仮想ＬＤＥＶ６１ｃ、ＬＤＥＶプール６８、及びＤＭＴ６４の構成例を示す。

【図５】ＰＭＴ６３の構成例を示す。

【図６】本実施形態に係るＬＵ－ＬＤＥＶ管理テーブル１６２ｂの構成例を示す。

【図７】チャネル制御部１１０Ａの構成例を示す。

【図８】ＬＵが設定される場合に行なわれる記憶制御サブシステム１０２での処理流れを示す。

【図９】ホスト端末２００Ａの処理フローの一例を示す。

【図１０】記憶制御システム管理テーブル７９１の一例を示す図。

【図１１】チャネル制御部１１０Ａが、ホスト端末２００から第２コマンドとしてリードキャパシティコマンドを受けた場合に行なう処理フローの一例を示す。

【図１２】本実施形態の第１の特徴的な処理に関わる処理流れを示す。

【図１３】本実施形態の第３の特徴的な処理に関わる構成及び処理流れを示す。

【図１４】本実施形態の第４の特徴的な処理に関わる処理流れを示す。

【図１５】本実施形態の第５の特徴的な処理に関わる処理流れを示す。

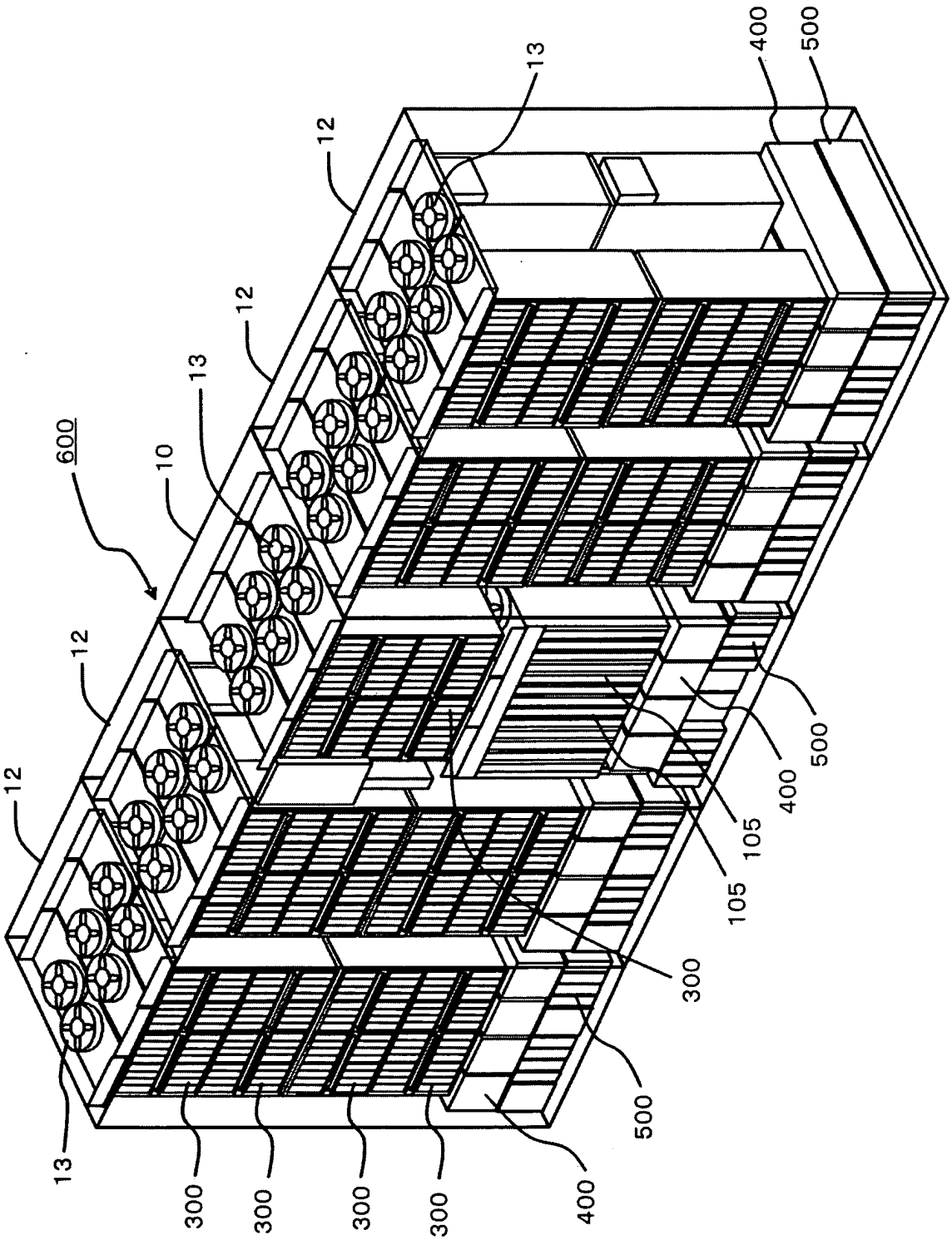
【図１６】本実施形態の第６の特徴的な処理に関わる処理流れを示す。

【符号の説明】

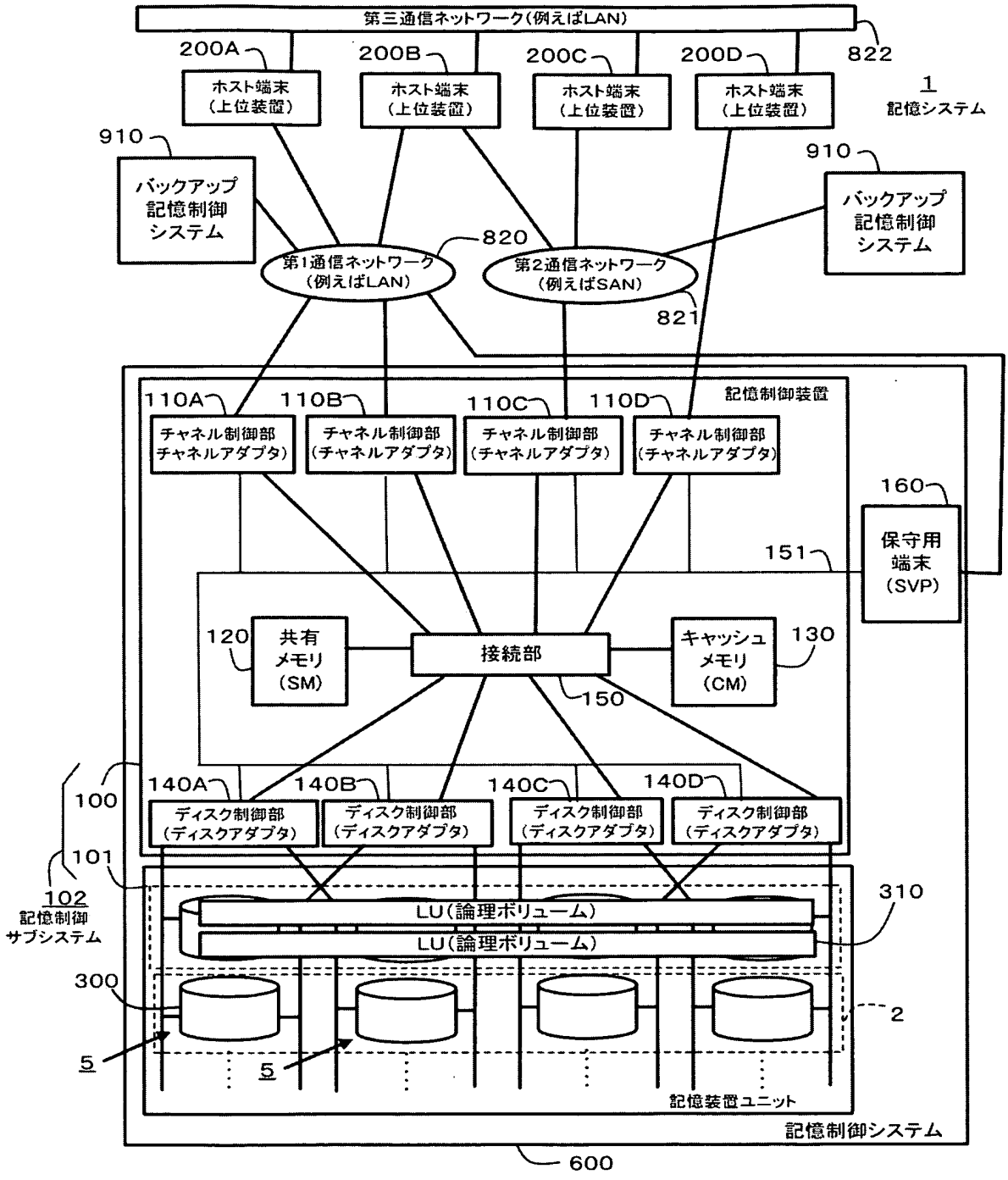
【０１４２】

１…記憶システム、２…ＲＡＩＤグループ、１０…基本筐体、１２…増設筐体、１００…記憶制御装置、１０１…記憶装置ユニット、１０２…記憶制御サブシステム、１１０Ａ～１１０Ｄ…チャネル制御部、１２０…共有メモリ、１３０…キャッシュメモリ、１４０Ａ～１４０Ｄ…ディスク制御部、１６０…保守用端末、２００Ａ～２００Ｄ…ホスト端末、３００…ディスク型記憶装置、３１０…論理ユニット、６００…記憶制御システム

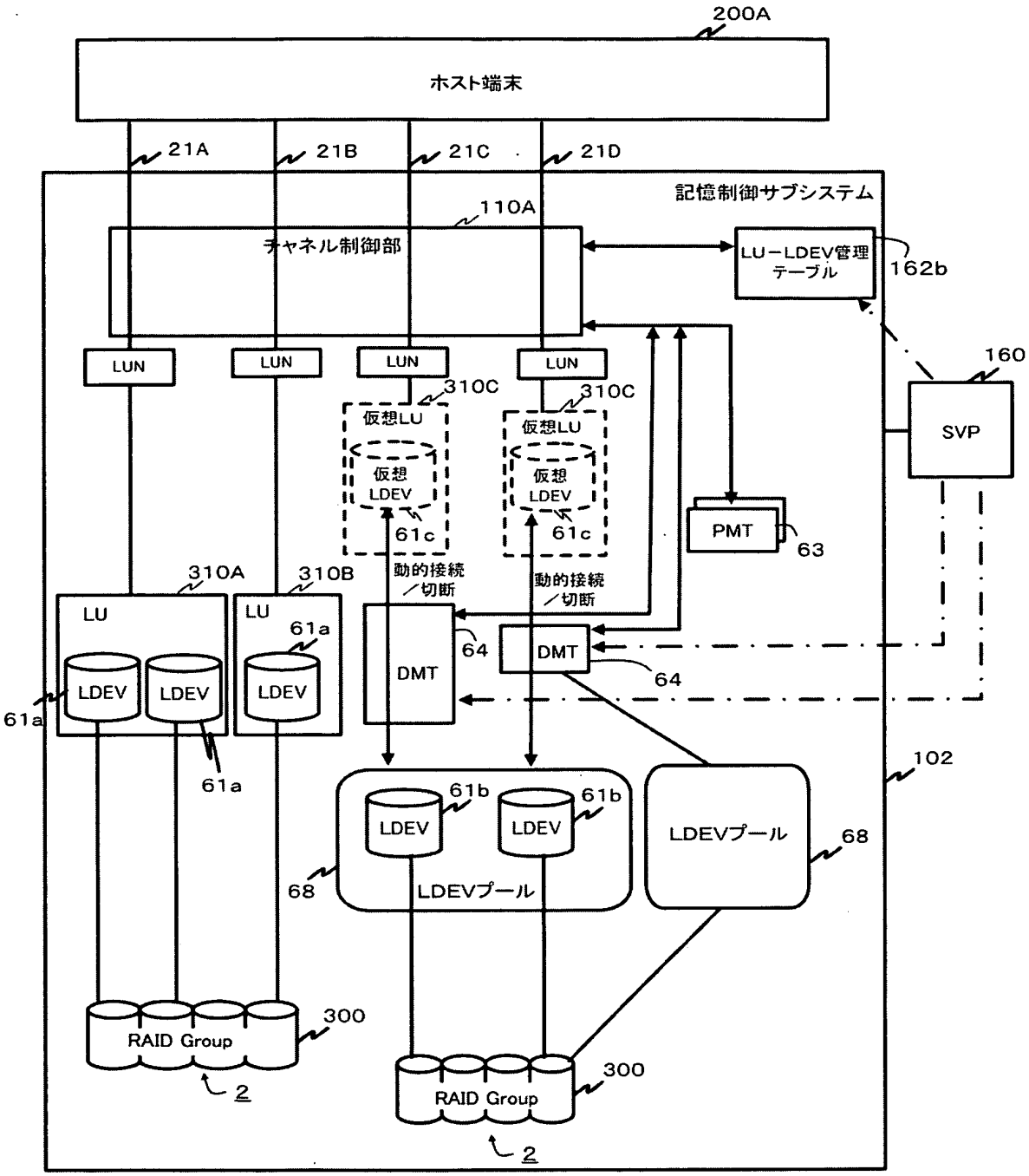
【書類名】図面
【図１】


400
500
13
12
12
13
600
10
12
12
500
400
105
105
300
500
13
300
300
300
300
400

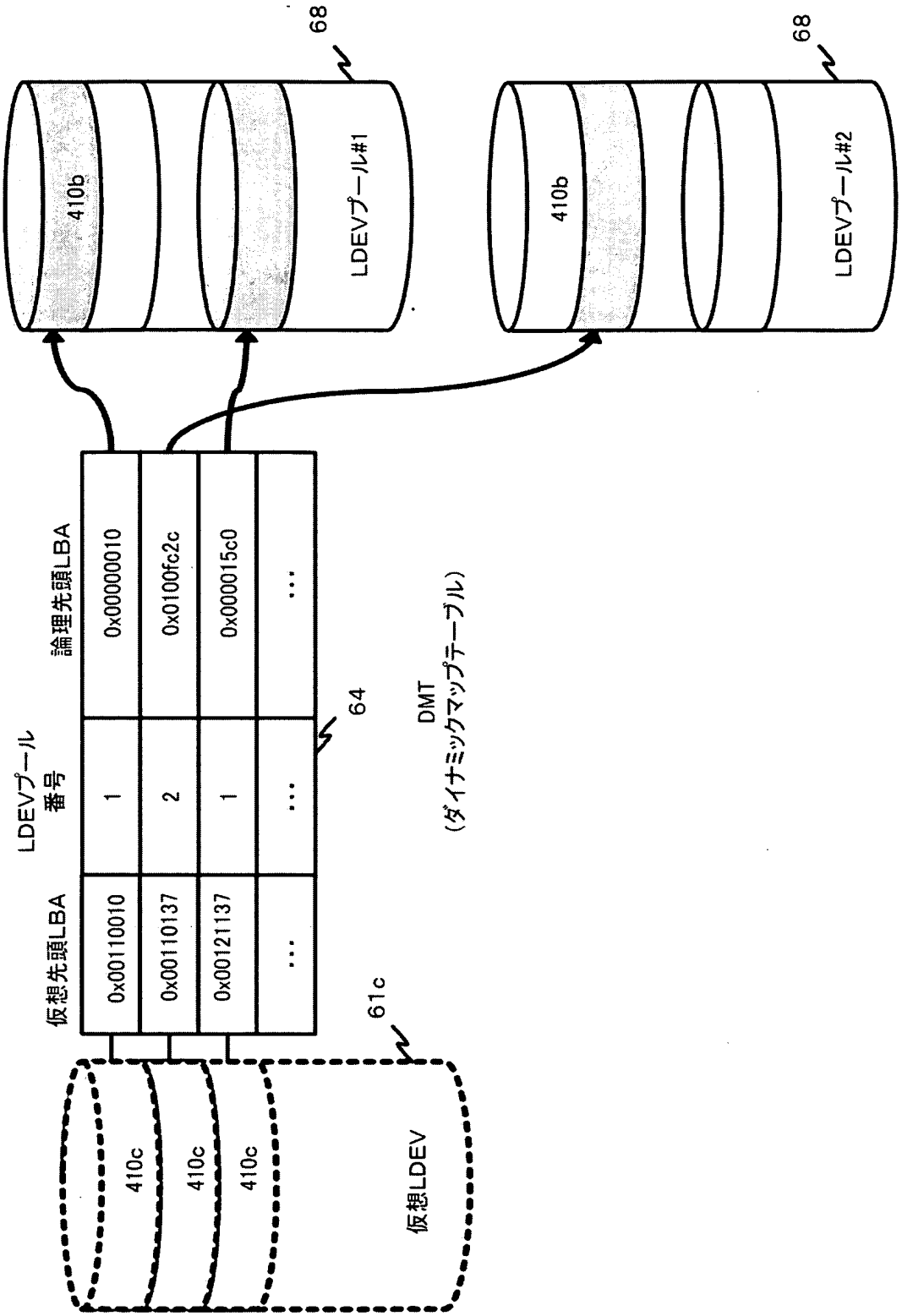
【図２】


第三通信ネットワーク(例えばLAN)
822
200A
200B
200C
200D
ホスト端末
(上位装置)
1
記憶システム
910
バックアップ
記憶制御
システム
820
第1通信ネットワーク
(例えばLAN)
第2通信ネットワーク
(例えばSAN)
821
記憶制御装置
110A
110B
110C
110D
チャネル制御部
(チャネルアダプタ)
160
保守用
端末
(SVP)
151
120
共有
メモリ
(SM)
接続部
キャッシュ
メモリ
(CM)
130
150
100
140A
140B
140C
140D
ディスク制御部
(ディスクアダプタ)
101
102
記憶制御
サブシステム
310
LU(論理ボリューム)
300
2
5
記憶装置ユニット
記憶制御システム
600

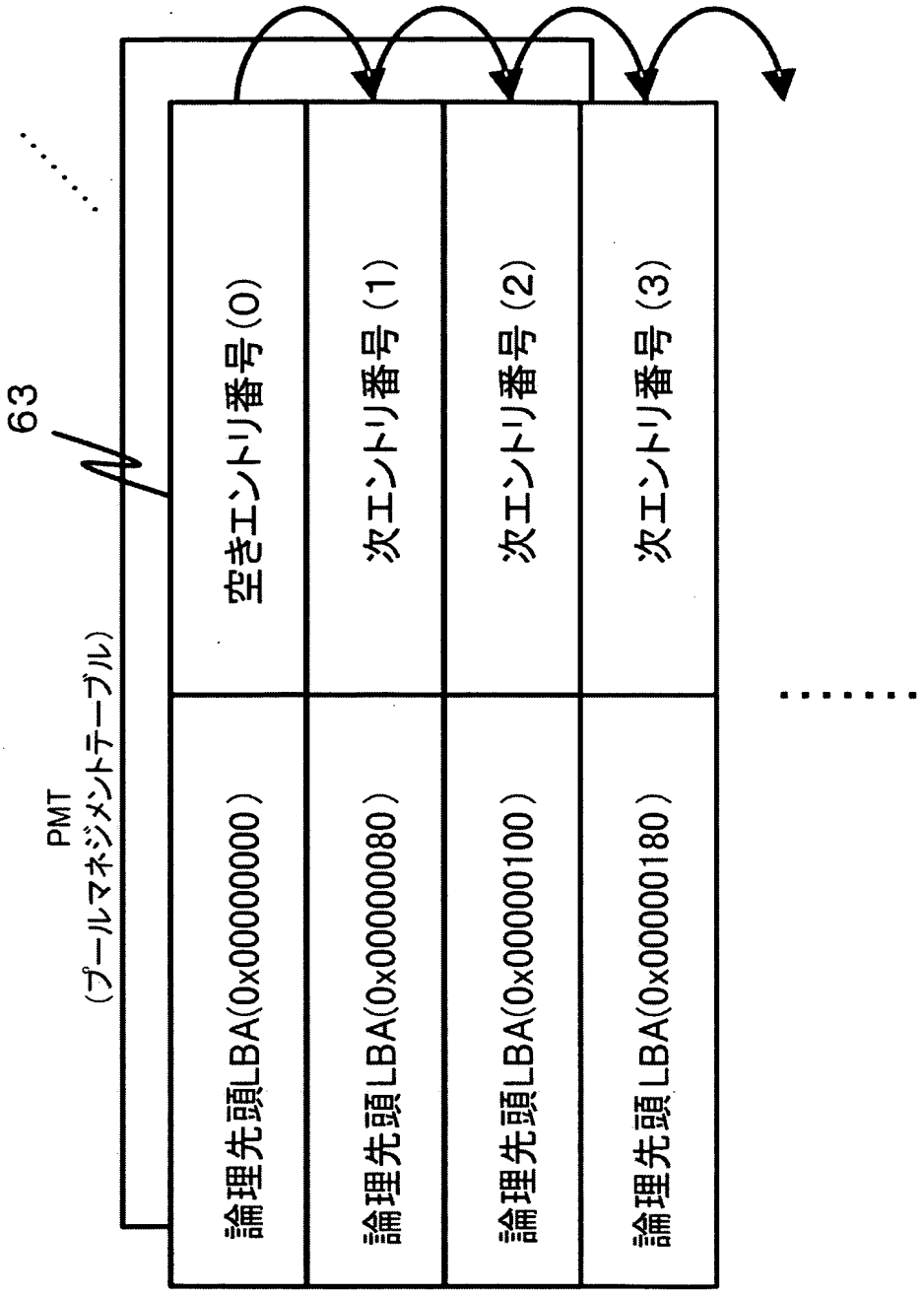
【図３】


200A
ホスト端末
21A
21B
21C
21D
記憶制御サブシステム
110A
チャネル制御部
LU－LDEV管理テーブル
162b
160
SVP
LUN
LUN
LUN
LUN
310C
310C
仮想LU
仮想LU
仮想LDEV
仮想LDEV
61c
61c
PMT
63
310A
310B
動的接続／切断
動的接続／切断
LU
LU
61a
LDEV
LDEV
LDEV
61a
61a
DMT
64
DMT
64
102
61b
61b
LDEV
LDEV
LDEVプール
68
68
LDEVプール
RAID Group
300
2
RAID Group
300
2

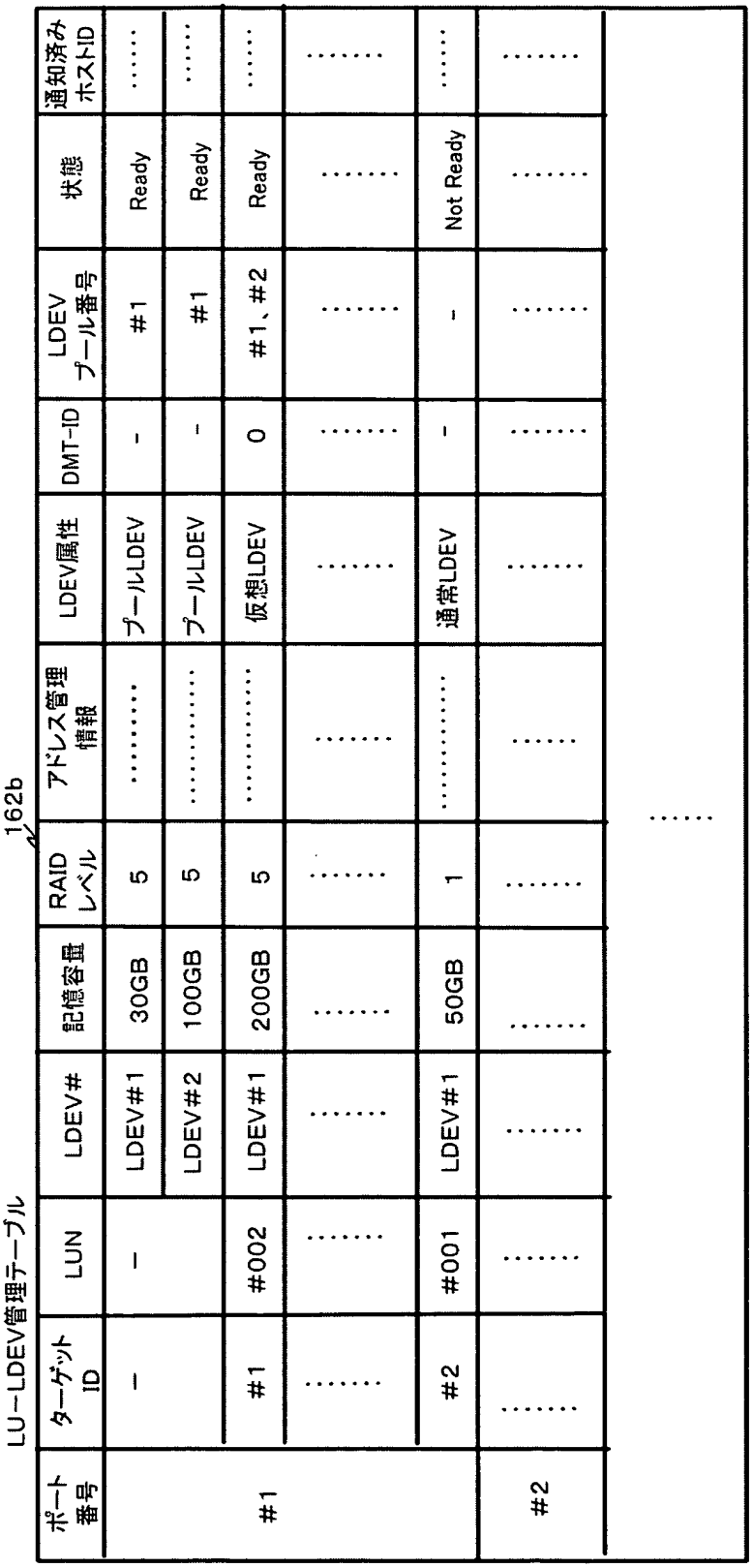
【図４】


68
68
410b
LDEVプール#1
410b
LDEVプール#2
論理先頭LBA
0x00000010
0x0100fc2c
0x0000015c0
...
LDEVプール
番号
1
2
1
...
64
DMT
(ダイナミックマップテーブル)
仮想先頭LBA
0x00110010
0x00110137
0x00121137
...
61c
410c
410c
410c
仮想LDEV

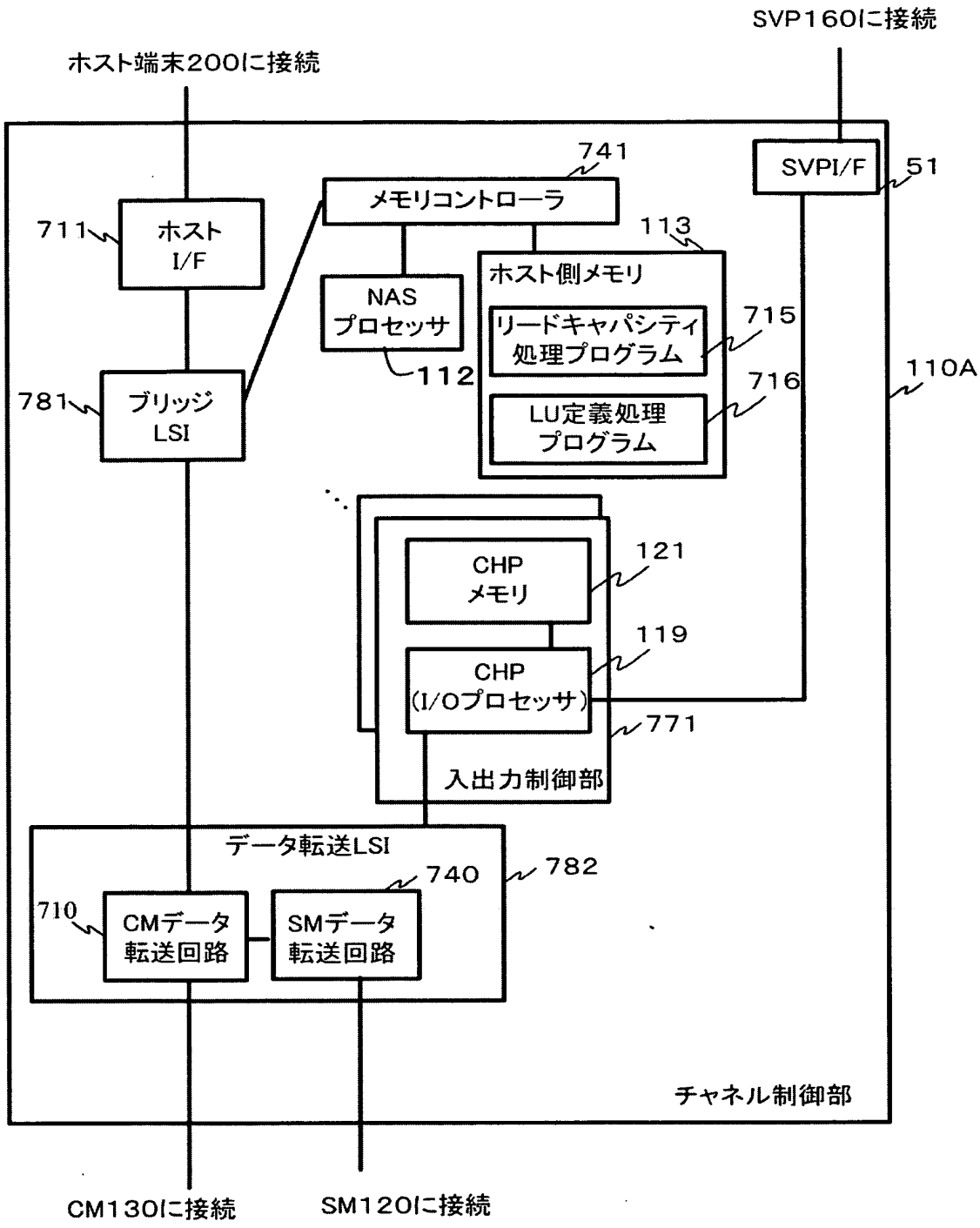
【図５】


63
PMT
(プールマネジメントテーブル)
空きエントリ番号(0)
次エントリ番号(1)
次エントリ番号(2)
次エントリ番号(3)
論理先頭LBA(0x00000000)
論理先頭LBA(0x00000080)
論理先頭LBA(0x00000100)
論理先頭LBA(0x00000180)

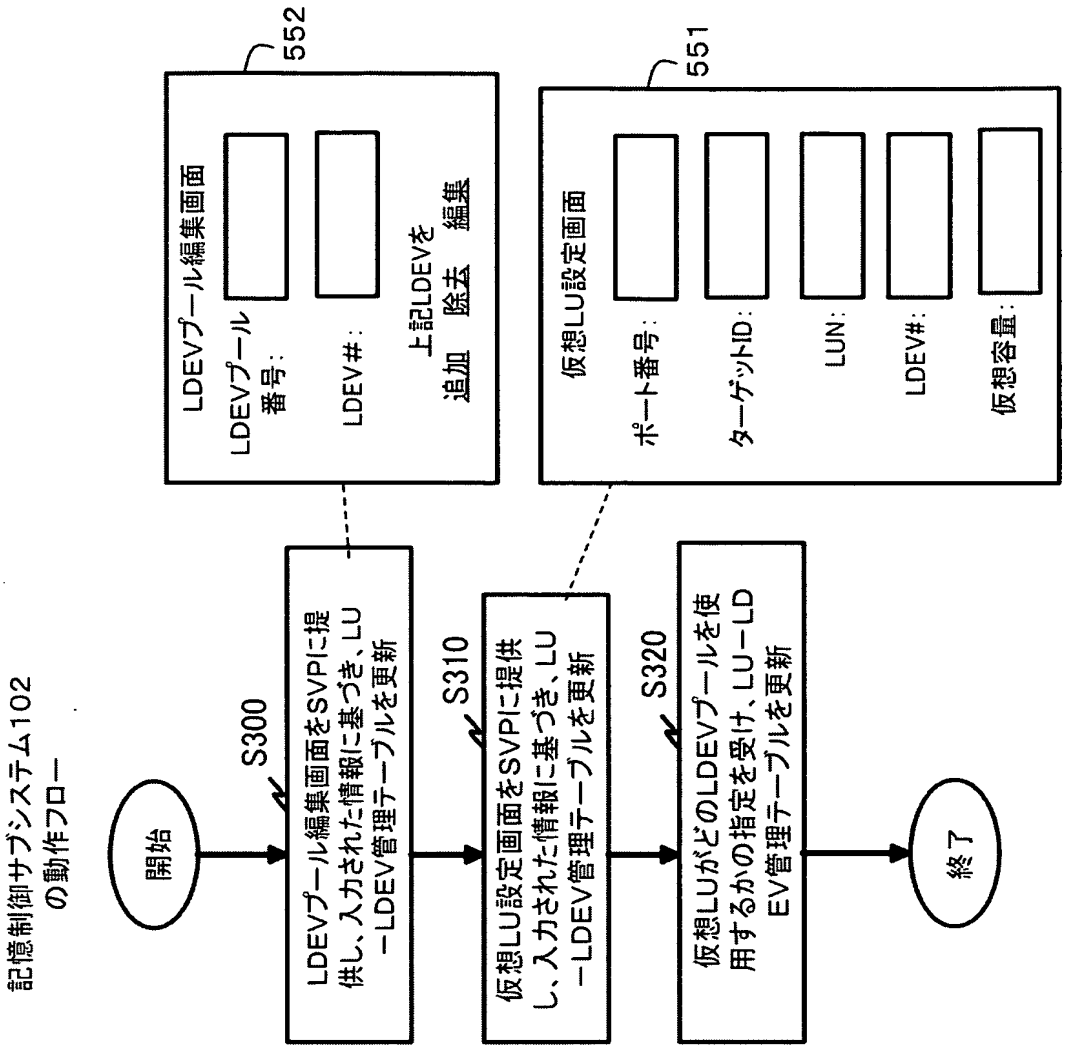
【図６】


LU－LDEV管理テーブル
162b
ポート番号
ターゲットID
LUN
LDEV#
記憶容量
RAIDレベル
アドレス管理情報
LDEV属性
DMT-ID
LDEV プール番号
状態
通知済みホストID
#1
－
－
LDEV#1
30GB
5
……
プールLDEV
－
#1
Ready
……
LDEV#2
100GB
5
……
プールLDEV
－
#1
Ready
……
#1
#002
LDEV#1
200GB
5
……
仮想LDEV
0
#1、#2
Ready
……
……
……
……
……
……
……
……
……
……
……
#2
#001
LDEV#1
50GB
1
……
通常LDEV
－
－
Not Ready
……
……
……
……
……
#2
……
……
……
……
……
……
……
……
……
……

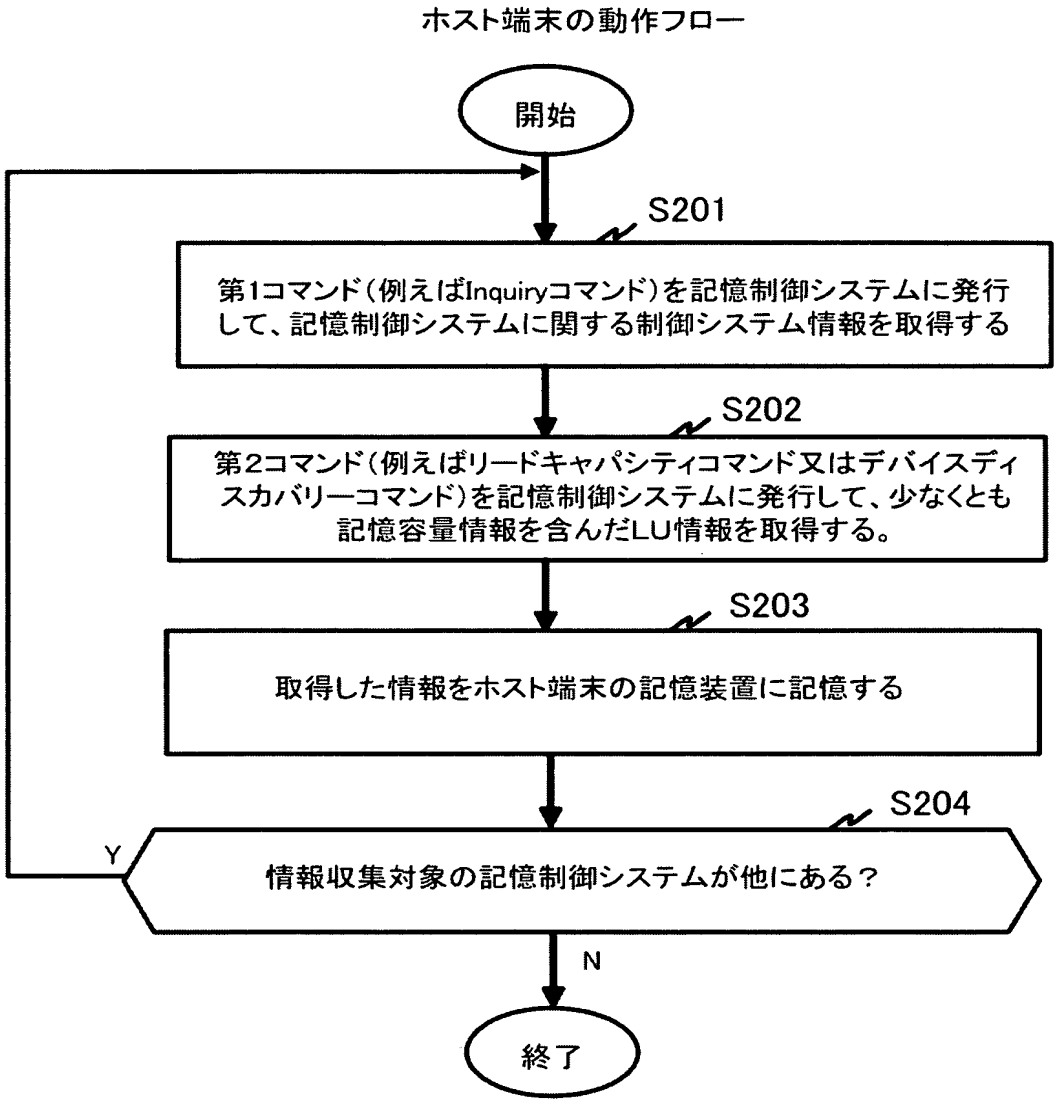
【図７】


SVP160に接続
ホスト端末200に接続
741
メモリコントローラ
SVPI/F
51
711
ホスト
I/F
113
ホスト側メモリ
NAS
プロセッサ
112
715
リードキャパシティ
処理プログラム
716
LU定義処理
プログラム
110A
781
ブリッジ
LSI
121
CHP
メモリ
119
CHP
(I/Oプロセッサ)
771
入出力制御部
データ転送LSI
740
782
710
CMデータ
転送回路
SMデータ
転送回路
チャネル制御部
CM130に接続
SM120に接続

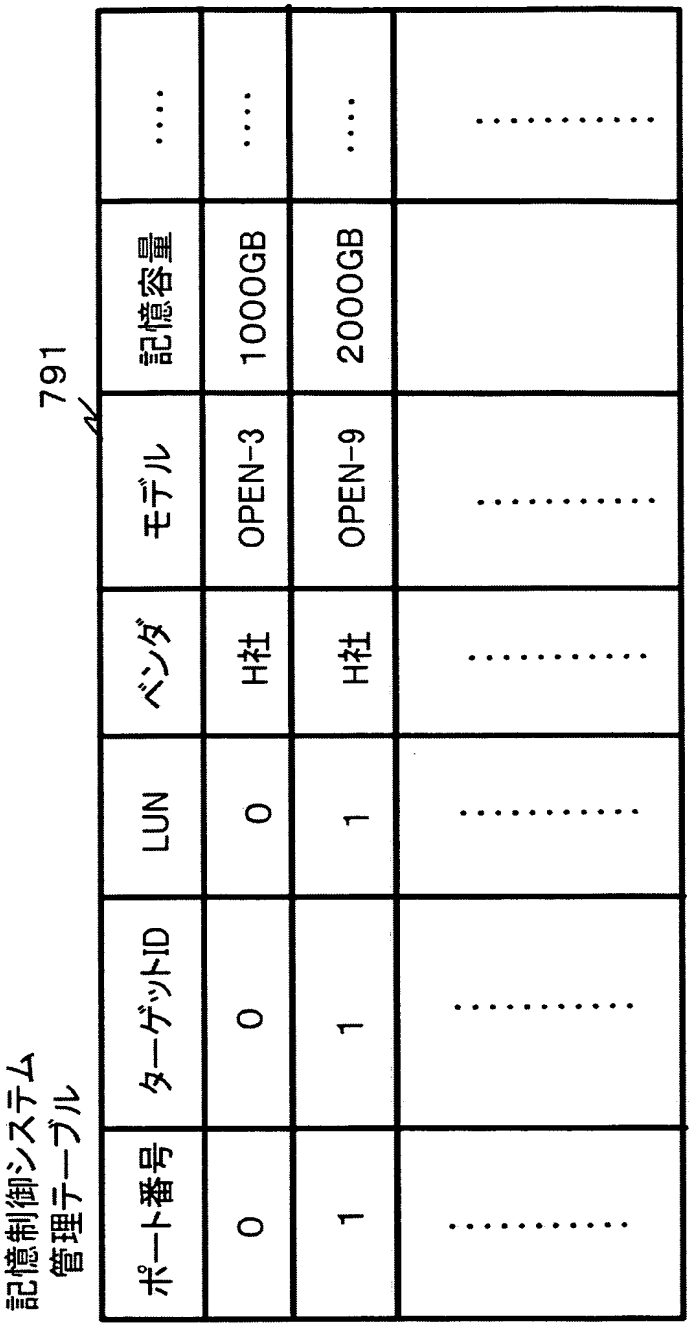
【図８】


記憶制御サブシステム102
の動作フロー
開始
S300
LDEVプール編集画面をSVPに提供し、入力された情報に基づき、LU－LDEV管理テーブルを更新
S310
仮想LU設定画面をSVPに提供し、入力された情報に基づき、LU－LDEV管理テーブルを更新
S320
仮想LUがどのLDEVプールを使用するかの指定を受け、LU－LDEV管理テーブルを更新
終了
552
LDEVプール編集画面
LDEVプール番号:
LDEV#:
上記LDEVを
追加　除去　編集
551
仮想LU設定画面
ポート番号:
ターゲットID:
LUN:
LDEV#:
仮想容量:

【図９】

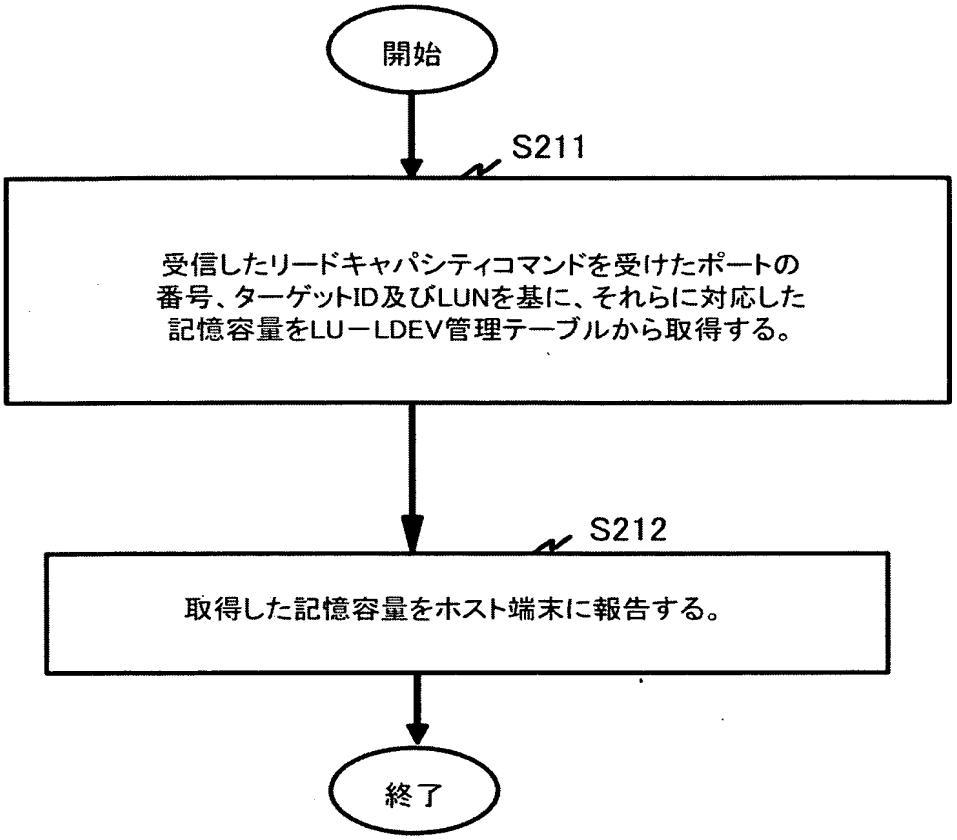

ホスト端末の動作フロー
開始
S201
第1コマンド（例えばInquiryコマンド）を記憶制御システムに発行して、記憶制御システムに関する制御システム情報を取得する
S202
第2コマンド（例えばリードキャパシティコマンド又はデバイスディスカバリーコマンド）を記憶制御システムに発行して、少なくとも記憶容量情報を含んだLU情報を取得する。
S203
取得した情報をホスト端末の記憶装置に記憶する
S204
Y
情報収集対象の記憶制御システムが他にある？
N
終了

【図１０】

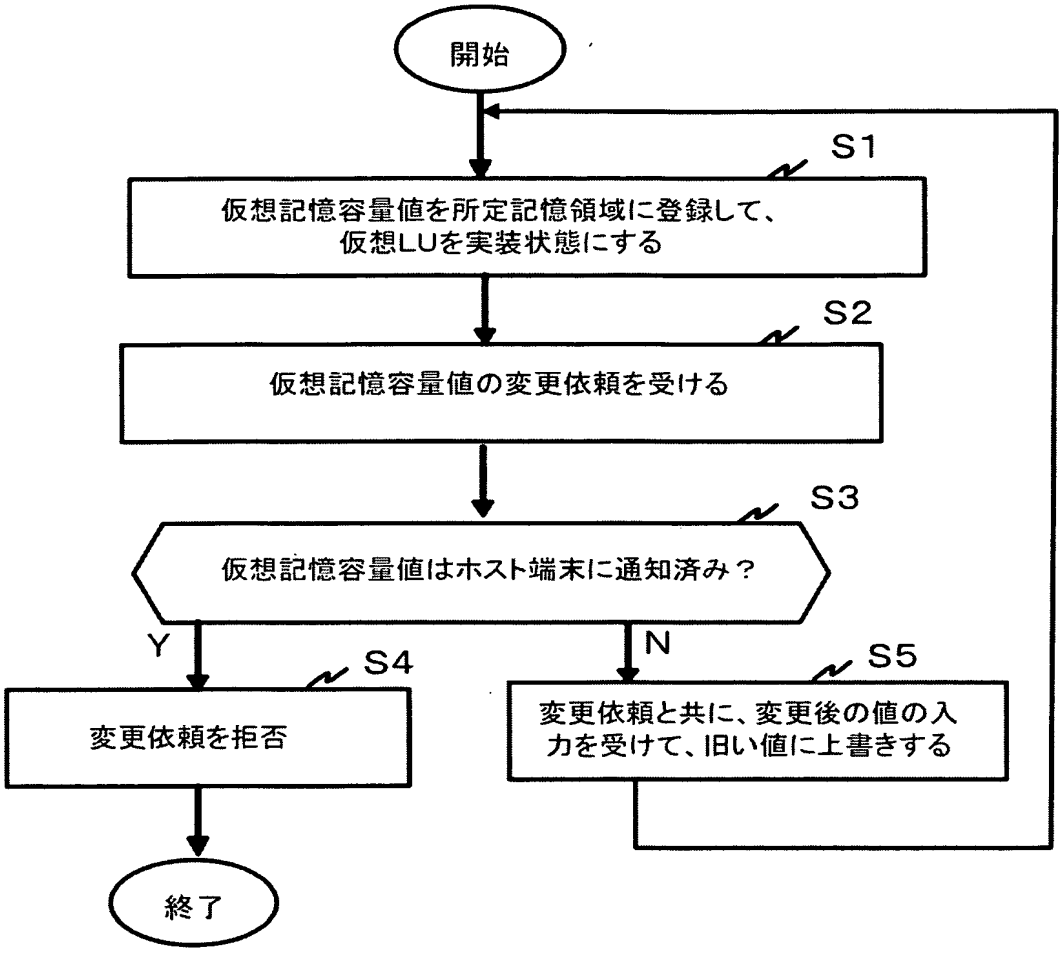

記憶制御システム
管理テーブル
791
ポート番号
ターゲットID
LUN
ベンダ
モデル
記憶容量
0
0
0
H社
OPEN-3
1000GB
1
1
1
H社
OPEN-9
2000GB

【図１１】

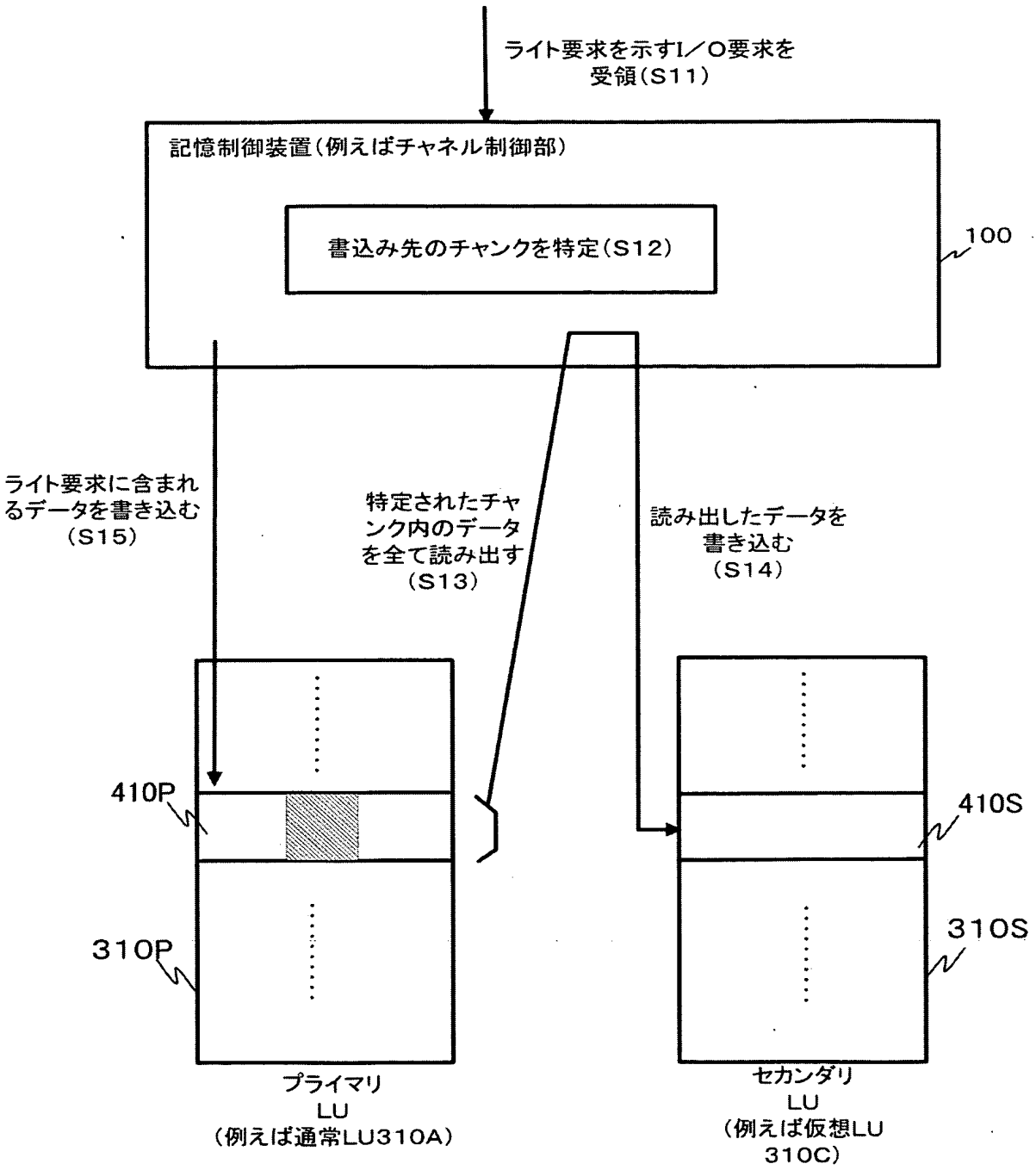
チャネル制御部の動作



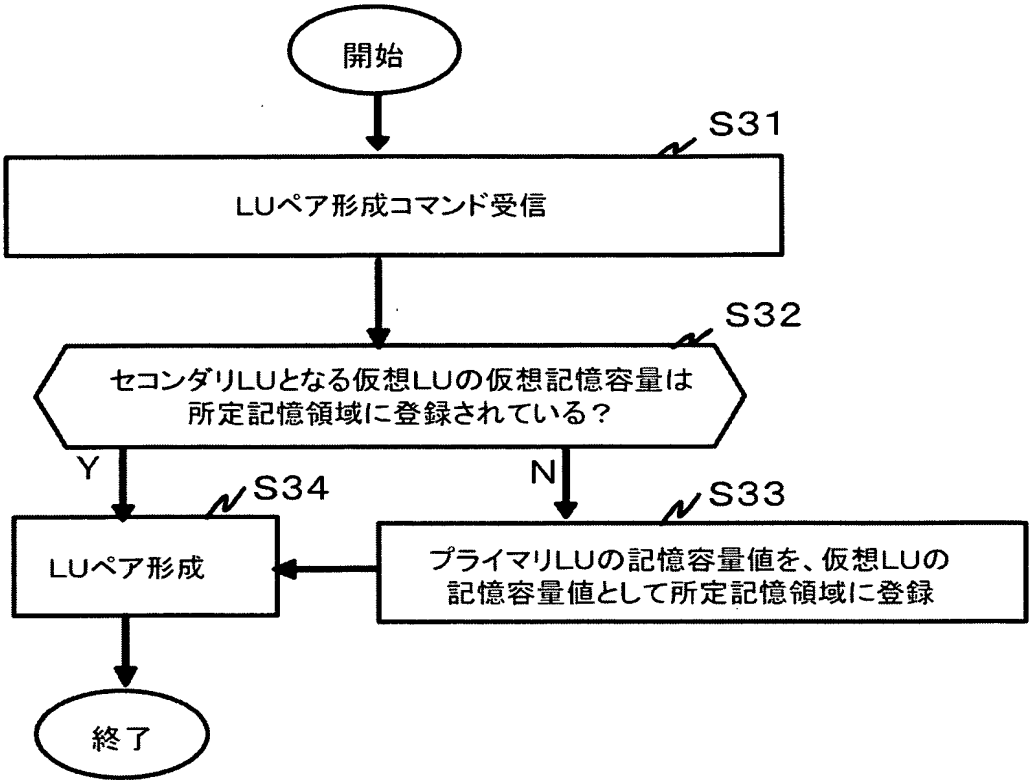
【図１２】


開始
S1
仮想記憶容量値を所定記憶領域に登録して、
仮想LUを実装状態にする
S2
仮想記憶容量値の変更依頼を受ける
S3
仮想記憶容量値はホスト端末に通知済み？
Y
N
S4
変更依頼を拒否
S5
変更依頼と共に、変更後の値の入
力を受けて、旧い値に上書きする
終了

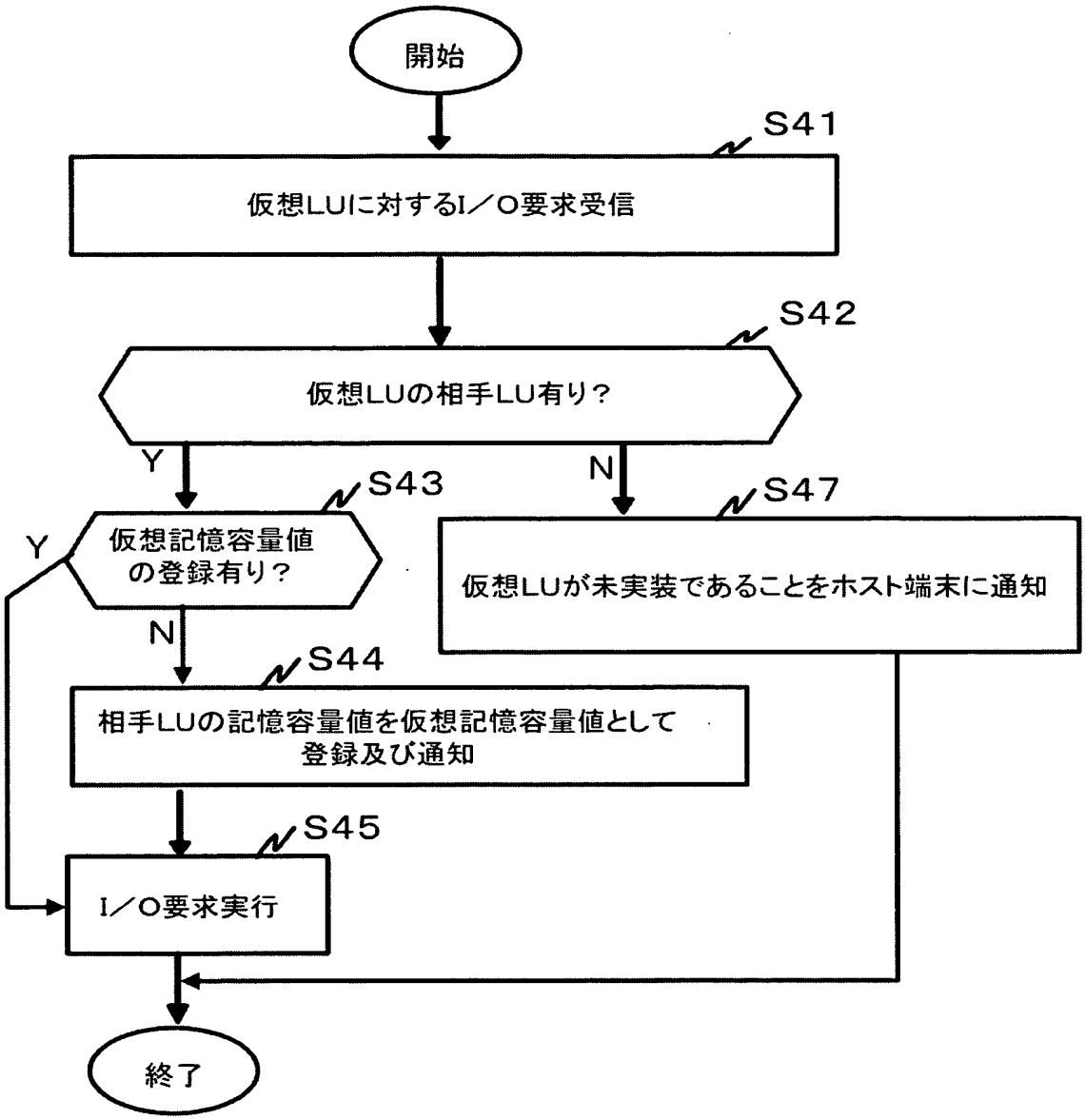
【図１３】


ライト要求を示すI／O要求を
受領(S11)
記憶制御装置(例えばチャネル制御部)
書込み先のチャンクを特定(S12)
100
ライト要求に含まれ
るデータを書き込む
(S15)
特定されたチャ
ンク内のデータ
を全て読み出す
(S13)
読み出したデータを
書き込む
(S14)
410P
310P
410S
310S
プライマリ
LU
(例えば通常LU310A)
セカンダリ
LU
(例えば仮想LU
310C)

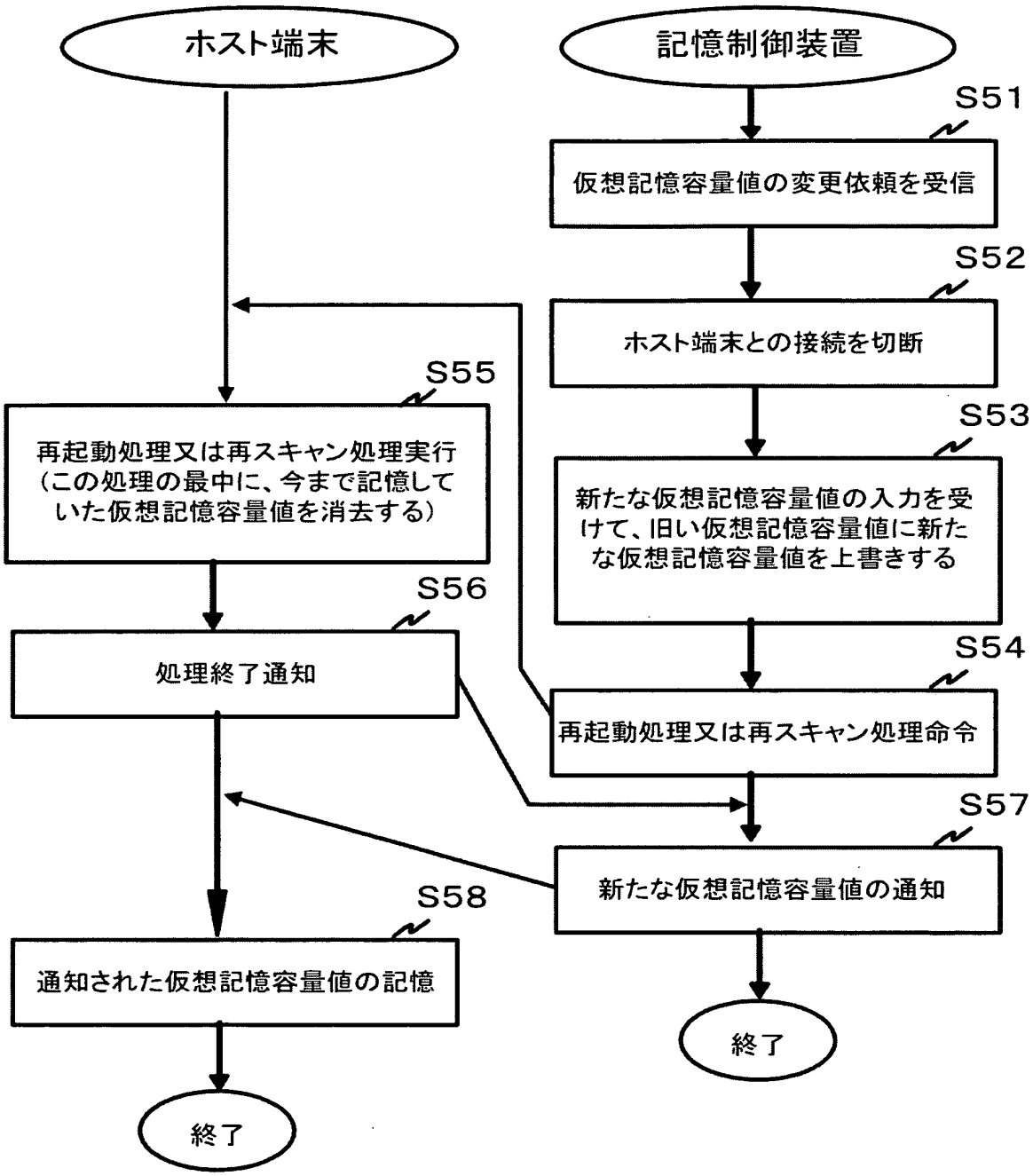
【図１４】


開始
S31
LUペア形成コマンド受信
S32
セコンダリLUとなる仮想LUの仮想記憶容量は
所定記憶領域に登録されている？
Y
N
S34
S33
LUペア形成
プライマリLUの記憶容量値を、仮想LUの
記憶容量値として所定記憶領域に登録
終了

【図１５】


開始
S41
仮想LUに対するI／O要求受信
S42
仮想LUの相手LU有り？
Y
N
S43
仮想記憶容量値
の登録有り？
Y
N
S47
仮想LUが未実装であることをホスト端末に通知
S44
相手LUの記憶容量値を仮想記憶容量値として
登録及び通知
S45
I／O要求実行
終了

【図１６】


ホスト端末
記憶制御装置
S51
仮想記憶容量値の変更依頼を受信
S52
ホスト端末との接続を切断
S53
新たな仮想記憶容量値の入力を受けて、旧い仮想記憶容量値に新たな仮想記憶容量値を上書きする
S54
再起動処理又は再スキャン処理命令
S55
再起動処理又は再スキャン処理実行（この処理の最中に、今まで記憶していた仮想記憶容量値を消去する）
S56
処理終了通知
S57
新たな仮想記憶容量値の通知
S58
通知された仮想記憶容量値の記憶
終了
終了

【書類名】要約書
【要約】
【課題】　ホスト端末に混乱を生じさせてしまう可能性を低減する記憶制御サブシステムが提供する。
【解決手段】　記憶制御サブシステム１０２に備えられる記憶制御装置１００は、共有メモリ１０２に記憶された仮想記憶容量値をホスト端末に通知し、そのホスト端末において仮想記憶容量値が記憶された後、その仮想記憶容量値を有する仮想記憶ユニットがホスト端末に接続されている間は、上記通知した仮想記憶容量値が変更されないようにする。
【選択図】図２

認定・付加情報

特許出願の番号	特願２００４－０３１１５０
受付番号	５０４００２０１２６０
書類名	特許願
担当官	第七担当上席　　００９６
作成日	平成１６年　２月　９日

＜認定情報・付加情報＞

【提出日】　平成16年 2月 6日

特願２００４－０３１１５０

出　願　人　履　歴　情　報

識別番号　［０００００５１０８］

１．変更年月日　１９９０年　８月３１日
［変更理由］　新規登録
住　所　東京都千代田区神田駿河台４丁目６番地
氏　名　株式会社日立製作所